THE UNIVERSITY OF ALBERTA

FORMAL OPERATIONS IN SECONDARY STUDENTS: A TEST OF THE IDEA OF INTELLECTUAL STRUCTURAL LIMITATIONS

by

Edward Desmond Hobbs

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

FALL, 1975

ABSTRACT

One application of the theory of formal operations is that absence of formal-operational structures in students entails restrictions on curriculum and instruction. The present study represents an empirical test of whether this idea of intellectual structural limitations applies to secondary school students (grades 7 through 12). Tests of volume conservation, controlling variables, and the combinatorial system were administered to a sample of grade 7 students and to a sample of grade 9 students, to test the hypothesis that a majority of secondary school students is concrete-operational. Tests were administered to groups of students and employed objective scoring criteria. Two one-hour instructional sequences, in controlling variables and in generating combinations, were administered to the grade 7 sample over a one-week period. The sample was retested two days after and one month after the instructional sequences to test the hypothesis that Substage III-B behavior on tests of controlling variables or generating combinations cannot be induced in a short period of time (seven days). Of the grade 9 sample, 80% were judged to display Substage III-A behavior and 50%

were judged to display Substage III-B behavior. Of the grade 7 sample, 51% were judged to display Substage III-A behavior and 47% were judged to display Substage III-B behavior on the pretest, while 64% were judged to display Substage III-B behavior on the posttest. No significant change occurred from posttest to retention test. The increase from pretest to posttest in the proportion of the grade 7 subjects classified as Substage III-B is significant ($X^2 = 3.71$, $df = 1$, $p < .05$). The increase from pretest to retention test in the proportion of the grade 7 subjects classified as Substage III-B is significant ($X^2 = 6.66$, $df = 1$, $p < .005$). The results may be taken to indicate that the idea of intellectual structural limitations should not be applied to the majority of secondary school students.

ACKNOWLEDGEMENTS

Assistance in the preparation of this thesis was given by Drs C. Brainerd, W. Brouwer, N. Hersom, S. Hunka, and A. Hughes.

The research could not have been carried out without the cooperation of Darrell Gates, Blair Hamilton, Don McPherson, and Greg Ross, Principals of Cornwallis, West Hants, Wolfville, and Central Kings High Schools in Nova Scotia. The assistance of John Chute, Bruce Gilbert, and Doug Mortimore in the testing and instruction was extremely valuable.

In particular I wish to acknowledge the continued and much appreciated assistance and encouragement given by Dr. Len Stewin, Dr. Marshall Nay, and especially Dr. Heidi Kass.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

CHAPTER I

INTRODUCTION

The present study is an attempt to test the validity of a proposition currently being advanced by some educators (e.g. Ball & Sayre, 1974; Renner et al, 1973; Lawson, 1973; Lawson & Renner, 1974; Wood, 1974). The proposition is that many or perhaps most secondary school students are not "formal operational" and are not able to learn certain concepts, nor to understand purely formal reasoning. For instance Renner et al (1973) claim that "no less than 77% of junior high students is concrete operational, and in order for these students to profit educationally and intellectually from school, they must (not should, must) be interacting with concrete objects, events, and situations." To quote further:

> This is particularly true at the tenth-grade level, where biology has traditionally been taught and most of that population enrolls in it. Any teacher facing a biology class can be nearly certain that about 70 percent of those present are concrete operational thinkers. Now consider such concepts as the DNA molecule, gene mutations, and the chemical basis of life. From our perspective, such topics and concrete operational learners are grossly incompatible [p. 347].

Ball and Sayre (1974) put it this way: "Quite bluntly, students may not be able to learn what you teach"

because they have not developed "the mental framework necessary for abstract thinking."

The idea that limitations in intellectual structures place limitations on what students may be able to learn or understand may be termed "the idea of intellectual structural limitations." This idea will be defined in more detail below, but the problem to which the present study is addressed is whether this idea applies to secondary school curriculum and instruction. (Secondary school is taken here to mean grades 7 through 12.)

The idea of intellectual structural limitations has been questioned before. Sullivan (1967, 1969) has cautioned that "the use of Piaget's stages as indicators of 'learning readiness' seems most premature and needs more careful consideration on both the research and theoretical levels [1967, p. 33]." Hamlyn (1971) says that to apply stage theory to education in this way may be "educationally disastrous," for there is the "necessity within education of trying to teach people what they cannot understand [p. 23]." Ausubel and Robinson (1969) see a danger that:

> . . . literal interpretation of the terms used to describe Piaget's stage of formal operations may be causing some educators to put an unnecessarily low ceiling on the complexity of intellectual tasks to which young children are subjected [p. 200].

While the above objections are based on conjecture, the present study represents an empirical test of the idea of intellectual structural limitations for the particular case of its applicability to secondary school curriculum and instruction.

A case will be made that the applicability of the idea of intellectual structural limitations to the secondary school rests on the validity of three statements:

(a) The onset of the stage of formal operations (Substage III-A) does not occur for a majority of students until the secondary school years or later.

(b) The full realization of formal operations (Substage III-B) is not attained for a majority of students during the secondary school years.

(c) Progression from Substage III-A to Substage III-B is necessarily slow (of the order of years) and is not subject to appreciable acceleration in the secondary school years.

It will then be argued that for structural limitations to be applicable to secondary school either (a) must be true, or else (b) and (c) must be true simultaneously; and that to establish the empirical falsity of (a) and at the same time either (b) or (c) would be to destroy the connection between the theory

of formal operations and the idea of structural limitations. The present research aims to do this.

The three statements will be cast into testable form by operationally defining the terms "attainment of Substage III-A," "attainment of Substage III-B," and "appreciable acceleration," so that ultimately the questions to which the present study seeks answers are threefold:

1. What is the incidence of students in secondary schools who fail tests (to be specified) designed to detect (a) recognition of the need to control variables, and (b) the acquisition of volume conservation?

2. What is the incidence of students in secondary school who fail tests (to be specified) requiring (a) experimental control of variables, and (b) generation of the set of all subsets from a set of four elements?

3. In an experiment with age and training as independent variables and acquisition of Substage III-B as criterion, are the following effects significant?

(a) training main effect;

(b) age-training interaction effect.

DEFINITION OF TERMS

The idea of intellectual structural limitations

A prerequisite to definition of the idea of intellectual structural limitations is to establish the meaning of the term structure. In Piagetian theory, thinking is the act of performing mental operations. These operations or mental actions are not isolated entities but are organized and coordinated in systems or structures. Stages of intellectual development are characterized by their structures: eight structures called *groupements* in the stage of concrete operations (Stage II in the theory); and two structures called the combinatorial system and the four-group in the stage of formal operations or Stage III (Piaget, 1950, p. 36; 1964a, pp. 8-9; 1964b; 1970; Piaget & Inhelder, 1968, p. 329). In Piagetian theory, the stage of formal operations is divided into Substages III-A and III-B, with Substage III-A representing the "genesis" of the structures and Substage III-B their "complete form" or "equilibrium," with a period of the order of three years normally intervening between the substages (Inhelder, 1970, pp. 21-26).

The idea of intellectual structural limitations then, is defined to consist of two notions. The first notion is

that if structures are absent in a given subject, then forms of reasoning which depend on these structures will have no significance for the subject. He will not use such forms of reasoning, nor understand them. The second notion is that certain concepts may be impossible for a subject to learn if he lacks the appropriate structures, because in Piagetian theory, subjects construct scientific concepts for themselves through interaction between their intellectual structures and empirical experience (Piaget, 1970).

<u>Substages III-A and III-B</u>

A given subject will be said to have <u>attained Substage III-A</u> if <u>either</u>

(a) he may be classified as a conserver of volume, <u>or</u>

(b) he gives evidence of seeing the need to control variables in an experiment involving two or more independent variables.

A subject will be said to have <u>attained Substage III-B</u> if <u>either</u>

(a) he demonstrates the ability to devise a controlled experiment involving two or more independent variables, <u>or</u>

(b) he is able to generate all members of the set of all possible subsets of a set of four elements.

These definitions will be justified in Chapter III, where precise operational definition of attainment of Substages III-A and III-B will be given in terms of the tasks administered.

Training and acceleration

The development by a subject of Stage III structures from genesis to complete form takes about three years "in a rich cultural environment [Inhelder, 1970, p. 25]." For the present purposes then, acceleration may be defined in relation to development within Stage III as any reduction below three years in the average time between Substages III-A and III-B.

The term appreciable acceleration is defined as the experimental induction in a group of Substage III-A subjects of Substage III-B performance on a given task in a period of the order of weeks. This definition also will be justified below.

The majority of a population is more than 50%.

DEFINITION OF THE PROBLEM

The central question is whether the idea of intellectual structural limitations is applicable to at least a majority of secondary students. If most secondary school students have not entered Substage III-A,

structural limitations could apply in a straightforward manner. The structures are absent and any learning which may rely on the structures is theoretically not possible. On the other hand, if most secondary school students have attained Substage III-A but not Substage III-B, then the idea of intellectual structural limitations might still apply, but only if a given curriculum item were seen to require full realization of formal operational structures for proper comprehension. In that case, two courses of action would be possible. The item could be postponed until Substage III-B, or an acceleration to Substage III-B in the particular content area could be undertaken (and this amounts to teaching the item). If acceleration over a matter of days were possible in a given area of curriculum content then there would be no necessary structural basis for postponing content.

Consequently, the relevance of the idea of intellectual structural limitations to secondary school rests on the validity of three statements:

(a) Substage III-A is not attained for a majority of students until the secondary school years or later.

(b) Substage III-B is not attained for a majority of students during the secondary school years.

(c) Progression from Substage III-A to Substage III-B in a given content area is necessarily slow (of the

order of years) and is not subject to appreciable acceleration.

The aim of the study is to establish the empirical falsity of (a) and at the same time either (b) or (c).

PROCEDURE

The study has two parts: a survey and a training experiment. In the survey, four tasks were administered to a sample of grade 7 students and a sample of grade 9 students. Two of the tasks were designed to test for Substage III-A and two were designed to test for Substage III-B. In the experiment, two one-hour instructional sequences were administered to the grade 7 sample in three 40-minute sessions over a period of one week. The two Substage III-B tasks and one of the Substage III-A tasks were administered two days later, and again one month later.

DELIMITATIONS

The problem is delimited to the application of the idea of structural limitations to the majority of secondary students.Neither the idea itself as an application of the theory of formal operations nor the theory of formal operations as such is under test. On

the other hand, the validities of these two (the idea and the theory) are not necessarily assumed.

The validity of application of the idea of structural limitations to elementary pupils or to special groups of secondary students (such as low ability students) is not under test.

A further delimitation is that neither the specific nature of the instructional sequence nor the relative efficacy of different instructional sequences is a prime consideration. The prime consideration in relation to the instructional sequence is whether Substage III-B behavior can be induced in secondary school students in a short period of time.

OUTLINE OF THE REPORT

In Chapter II, the literature related to age norms for the stage of formal operations, and training experiments in formal operations is reviewed. In Chapter III, the tests are described, their validity discussed, and the method of their administration described. The data are presented, analyzed and discussed in Chapter IV. Chapter V contains a summary with conclusion and recommendations for further research.

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter has two sections. The first section is a review of a number of reports of research which include the administration of formal-operational tasks. A very wide variation exists in the proportions of subjects classified by the various experimenters as being in the period of formal operations. In the review, this variation is seen to be largely the result of wide variation in the criteria used by the various experimenters to classify subjects as formal operational. The second section of the chapter is a review of a number of experiments involving training in formal operations, concluding with a brief discussion of some theoretical problems associated with training experiments.

AGE NORMS FOR FORMAL OPERATIONS

The Genevan group consistently finds a majority of subjects in Substage III-A by the age of 11 or 12, and in Substage III-B by the age of 14 or 15. These researchers do not often give detailed descriptions of the samples studied, but apparently the approximate ages for the substages given in *The growth of logical thinking*

(Inhelder & Piaget, 1958) were based on "a somewhat privileged population [Piaget, 1972]." The method used in these experiments is for a trained experimenter to pose a problem to the subject and to observe the subject as he attempts to solve it. The experimenter takes note of all the subject says and does; not confining his activity to observing but actively questioning the subject. On the basis of this observation, the experimenter makes a subjective judgment as to the stage level of the subjects.

Other experimenters who have attempted to replicate or extend the Genevan experiments with other populations have not generally confirmed these age norms, and have found variation in age norms according to the particular task used. In reviewing these experiments it will be convenient to adopt a special designation for tasks based on those in _The growth of logical thinking_. These tasks will be represented by the letters GLT followed by the chapter number in which the task is described. For example, GLT-4 represents a task based on the pendulum problem in Chapter 4 of _The growth of logical thinking_.

Higgins-Trenk and Gaites (1971) tested 162 subjects (ages 13 to 18) in a Wisconsin high school. They

classified 43% as formal operational on a volume conservation task, and 32% as formal operational on "open-ended situational dilemmas." Higgins-Trenk and Gaites concluded that "the normal adolescent is unlikely to reach the level of formal operational thought until his late teens or early twenties if he reaches it at all, and these results suggest he may well not."

Renner and Stafford (1973) tested 588 students in grades 7 through 12 in Oklahoma secondary schools. In grade 7, 13% were classified as "post-concrete operational," and a further 3% as formal operational. In grade 12, 15% were classified as "post-concrete operational," and a further 18% as formal operational. Scores were allotted to subjects according to their stage level on each task as follows: Substage II-A - 1 point, Substage II-B - 2 points, Substage III-A - 3 points, and Substage III-B - 4 points. Scores for individual tasks were added and subjects scoring 12 to 14 points were classified as "post concrete," while subjects scoring 15 or 16 points (the possible) were classified as "formal operational." The tasks were administered by means of individual interviews, but only 20 minutes were given for the six tasks. Subjects must have had little time to think about their responses, and experimenters must have been forced to make snap

judgments. The criterion of perfect or near-perfect scores for classification as "formal operational" is a very rigorous one.

Friot (1970) tested 210 grade 8 and 9 students in an Oklahoma junior high school at the beginning and end of a school year. Subjects were classified as being in the stage of concrete operations, in a "transitional stage," or in the stage of formal operations. On the pretest, 18% were classified as being beyond concrete operational (that is either transitional or formal), and on the posttest,56%. Variation according to which science course was studied was found in the posttest. Like Renner and Stafford, Friot used a battery of six tests: conservation of quantity, conservation of volume, GLT-1, GLT-2, GLT-3, and GLT-4. Points were allotted in the fashion of the Renner and Stafford study. Further inspection of Friot's results shows that 66% of the grade 8 subjects gave formal-operational responses to at least one of the tasks on the pretest. This is over three times the proportion classified by Friot as transitional or formal.

Lawson (1973) classified 84% of 134 high school biology, chemistry, and physics students in Oklahoma as being beyond Stage II: 28% as "postconcrete," 38%

as Substage III-A, 14% as "transitional formal," and 4% as Substage III-B. Lawson also used a battery of six tests and a points scoring system.

Tomlinson-Keasey (1972) tested 30 girls, mean age 11.9, from a California school and 30 women, mean age 19.7, from Trenton State College, New Jersey. Thirty-three per cent of the girls were classified in Stage III: 29% in Substage III-A and 4% in Substage III-B; while 67% of the college women were classified in Stage III: 41% in Substage III-A and 26% in Substage III-B. The tasks used were GLT-3, -4, and -11. Individual interviews were used and subjective judgments as to stage level were made with the aid of sets of "behaviors characteristic of each stage of development." Tomlinson-Keasey reported "ambiguities and difficulties involved in judging each individual's protocol." A training session followed the three tasks, and the results of the posttests are reviewed in the second section of this chapter.

Wheeler and Kass (1974) tested high school students in Alberta. They classified 72% of 99 grade 12 chemistry students in Stage III: 61% in Substage III-A and 11% in Substage III-B. Two versions of GLT-7 were administered

to groups of subjects in written form.

Jackson (1965) tested normal and subnormal children in England using five tasks: GLT-2, -4, -8, -9, -11, and -12. Forty-eight normal subjects were tested in each of the age groups, 11, 13, and 15 years old. Of the 11-year-old subjects, 6% were classified as Substage III-A (none as Substage III-B). Of the 13-year-old subjects, 42% were classified as Stage III: 38% as Substage III-A and 4% as Substage III-B. Of the 15-year-old subjects, 48% were classified as Stage III: 38% as Substage III-A and 10% as Substage III-B. Variation was found according to task administered, the proportions of subjects classified as Stage III ranging from 6% on GLT-2 to 33% on GLT-4.

Stephens _et_ _al_ (1971) administered a battery of 29 Piagetian tasks to 75 normals and 75 retardates as part of a longtitudinal study. The mental ages at which 66% of subjects succeeded on the tasks were determined. Five different formal abilities (related to volume conservation) were attained by normals (IQ 90-110) according to the 66% criterion at mental ages 11, 13, 16, 17, and 24.

Bredderman (1973b) tested 27 grade 5 and grade 6 American students at three one-month intervals in a training experiment. In the first test, the proportion of students in stages II, III-A, and III-B were respectively 41%, 55%, and 4%; and in the third test, 8%, 44%, and 48%.

Bredderman used a test of controlling variables (GLT-3) in which the experimenter set the problem to a single subject by means of a set of prepared questions. Bredderman's experiment will be considered again below, under the heading of training experiments.

One indicator of formal thought which has been used by a number of experimenters is volume conservation, which is supposed to be an early acquisition in Stage III (Inhelder & Piaget, 1958, p. 36). Studies involving volume conservation have been reviewed elsewhere (Hobbs, 1972). Reported proportions of 11-year-old subjects classified as volume conservers vary from 41% (Lovell & Ogilvie, 1961) to 80% (Inhelder & Piaget, 1958).

In a review of literature related to the period of formal operations, Farrell (1969) concluded that while experiments at Geneva had shown that formal operations "begin to emerge" at the "age of about 11," other research indicates "some flexibility as to the age at which formal operations begin to develop and that at which the structure is in equilibrium." Farrell noted the "possible effect" of cultural differences (including schooling) on the age norms.

In a more recent review, Howe (1974) concluded that the results of the experiments by Renner and Stafford (1972), Stephens _et al_ (1971), Lovell (1961; 1968),

> . . . and other studies make it clear that only a minority of adolescents seem capable of, or find it necessary to, use this mode of thought [formal operations]. A few are able to solve Stage III A problems as early as 11 or 12, but most cannot solve truly formal operational problems until late in the teens or later, if at all.

Howe's conclusions may be a little premature in view of the variation in percentages of 13-year-old subjects (grade 7 level) classified as Stage III by various experimenters:

16%	---	Renner & Stafford (1973)
36%	---	Elkind (1961c)
42%	---	Jackson (1965)
60%	---	Hale (1973)
66%	---	Stephens _et al_ (1971)
75%	---	Somerville (1974)

Any comments on the wide disparity in these results would be superfluous. A weak explanation of the variation is that the proportions are of different populations in different cultural milieux (all populations are English speaking, middle-class, normal children). The disparity is more likely to be due to the variation in criteria for stage classification used by the various experimenters. Part of the variation in criteria is in the choice of task. Piagetian theory gives no basis for comparison of results obtained on different tasks which are supposed to test for the same structure (Piaget, 1971).

Another part of the variation in criteria is in deciding just what behavior is to constitute guidelines for experimenters' judgment on a given subject. An indication of the complexity of this problem is provided by Somerville (1974). In Somerville's study a sample of 236 Australian children from 10 to 14 years old was given the pendulum problem (GLT-4). The experimenter's aim was to develop a "detailed scoring system for performance." Children were tested individually, and "presentation of the task and the method of questioning were based as closely as possible on the work of the original authors [Inhelder & Piaget]." Subjects were rated on a three point scale on each of 12 aspects of their performance. In accordance with the "patterns of performance ratings" found, nine stages of development were defined, two corresponding to Substage III-B, three corresponding to Substage III-A, and four corresponding to Stage II. Of the 13-year-old subjects, 75% were classified as Stage III including 33% as Substage III-A and 42% as Substage III-A. Somerville's complex scoring system is indicative of the wide variety of behavior patterns which may be found at a given stage level on a given task.

TRAINING STUDIES

Four training experiments in formal operations have been reviewed by Howe (1974): Fishbein _et al_ (1970b) successfully trained most of 60 subjects, some as young as 10 years, in the combinatorial system. They claimed "structural acquisition" by their subjects, that is, learned ability which generalized to a different problem. Tomlinson-Keasey (1972) trained 72 subjects in GLT-3, -4, and -11 in a 45-minute session. Gains were found in stage level on an immediate posttest consisting of the same tasks, but these gains did not persist or transfer according to a test one week later using an altered version of GLT-3 and versions of GLT-7 and -12. Brainerd and Allen (1971) successfully trained "density conservation" to 52 grade 5 students in a single session. Transfer to volume conservation was found. Seigler _et al_ (1973) taught 10- and 11-year-old subjects (N = 24) to solve the pendulum problem (GLT-4). After a single training session "the 10- and 11-year-olds were able to exercise formal operations logic and to produce solutions closely resembling those cited by Inhelder and Piaget as exemplifying the highest stage of reasoning."

In a review of training in both concrete and formal operations Beilin (1971) expressed surprise at "how

little research is addressed to" the training of formal operations. He found the experiments reviewed to be unsuccessful in inducing skills related to combinatorial thinking (Ennis *et al*, 1969; Anderson, 1965), or unconvincing because of absence of a test of transfer (Fishbein *et al*, 1970a).

Successful training experiments have been reported by Case and Fry (1973) and Bredderman (1973b). Case and Fry trained a "group of 15 low SES 14-year-olds who had not yet reached Piaget's stage of 'formal operations'" in designing controlled experiments and criticizing poorly controlled experiments. After 12 weeks the experimental group performed significantly better (t = 8.4, df = 28, p < .001) than a matched control group on a test related to controlling variables. Since controlling variables is a formal-operational scheme, this may be considered to be a successful training experiment in formal operations. Bredderman trained 27 grade 5 and grade 6 students in controlling variables. Four training sessions were used, spaced over one month. The design included a pretest, a posttest, and a retention test one month after the posttest. Bredderman concluded that in the posttest "nearly half of the students having mean age 11.8 years could be classified in late formal operations with regard to their ability to control variables."

The overall picture seems to be that acceleration

of formal-operational behavior may be not only possible, but also relatively easy, although the evidence remains quite thin.

A methodological problem related to the training of operational structures has been raised by Inhelder and her co-workers (Inhelder et al, 1966; Inhelder & Sinclair, 1969). They claim that attempts to train subjects in conservation tasks may succeed in inducing subjects to respond correctly without their acquiring "truly logical structures [Inhelder & Sinclair, 1969, p. 19]." They recommend rigorous criteria for acquisition of logical structures to eliminate responses which "are examples of either a specific strategy learned in one type of problem or of a limited operational structure, but not of the acquisition of a true operational structure." The specific criteria they recommend are summarized as follows by Kuhn (1974): "To be considered as having genuinely attained conservation, a trained subject must make responses that (a) include appropriate explanations, (b) persist over time, and (c) generalize to nontrained material."

With regard to the first criterion (requirement of verbal explanations), Brainerd (1973) has shown that such a requirement makes tests subject to systematic Type II error. He argues that in Piagetian theory, cognitive

structures do not depend on verbal behavior, therefore verbal explanations constitute sufficient but not necessary conditions for the presence of cognitive structures. Although Kuhn questions Brainerd's argument, she shows that in posttests of training experiments, explanations cannot be considered to be even sufficient conditions of structure. Verbal explanations are not required in the tasks used in the present study.

The second and third of the three criteria for presence of structure (retention and transfer) are employed in the present study. There are theoretical difficulties associated with both of these. One difficulty relates to the extent of retention or transfer which may be acceptable as sufficient evidence of "genuine" structural acquisition. Kuhn (1974) has shown the ambiguity in the retention criterion: that structural acquisition is not a necessary condition for retention. The transfer criterion would be more convincing if there were never or seldom any variation in subjects' behavior on different tasks supposedly related to the same structure. In any case, in the experiment described in Chapter III, subjects were not taught how to do the specific tasks used in the posttest, so that transfer was required, and a retention test was administered one month after the posttest.

CHAPTER III

PROCEDURE

In Chapter I, a case was made that the application of the theory of intellectual structural limitations rests on the validity of the following three statements:

(a) Substage III-A is not attained for a majority of students until the secondary school years or later.

(b) Substage III-B is not attained for a majority of students during the secondary school years.

(c) Progression from Substage III-A to Substage III-B in a given content area is necessarily slow (of the order of years) and is not subject to appreciable acceleration.

For structural limitations to be applicable to secondary school, either (a) must be true, or (b) and (c) must be true simultaneously. The procedure which follows is designed to establish the empirical falsity of (a) and at the same time either (b) or (c).

Consequently, the study has two aims: to determine the proportions of secondary school students who may be classified as Substage III-A or Substage III-B; and to test whether appreciable acceleration can be induced in the progression from Substage III-A to Substage III-B.

Consequently, the study has two parts: a survey and an experiment. Both involve the testing of subjects in order to classify them according to substage. The tests used for this purpose, together with the criteria used for classification, represent operational definition of the terms "attainment of Substage III-A" and "attainment of Substage III-B."

The tests will now be described and the validity of the operational definitions they represent will be discussed. The procedures used in the survey and experiment will then be described.

TESTS

Tests for formal operational structures usually follow procedures described in The growth of logical thinking (Inhelder & Piaget, 1958). In these procedures, a single task is used to classify subjects somewhere on the full sequence of development. For instance a given subject might be classified as Substage III-B, Substage III-A, or Substage II on the basis of his behavior on the one task. Separate tasks are not used for each substage. The experimenter's subjective judgment usually enters into this multiple classification process.

In the present study, an attempt is made to provide a basis for classification of subjects which

requires minimum judgment by the experimenter once the criterion has been decided for a particular substage. In order to do this, each task was designed to test for one substage only. Two tasks were used to test for Substage III-A and two others to test for Substage III-B.

The wide disparity in the reported percentages of Stage III subjects at age 13 years given in Chapter II was seen to be probably a function of differing criteria for classification. The criterion problem--deciding what will be the criteria for classification of subjects according to stage level--is endemic to research related to stage or structure theories, because stage level or structure can never be determined directly but always has to be inferred from observed behavior on tasks. The problem of establishing criteria for a stage level or structure has two aspects: deciding on the form of the task, and deciding on what behavior elicited by the task will constitute evidence for the inferred stage level or structure.

These two problems will now be discussed in general terms, starting with the latter problem. The specific tasks used in the study will then be described.

Inference criteria

For classification into Substage III-B, most researchers seem to require a subject to succeed completely on a task, or even on each of a number of tasks. To do this is to assume that what a given subject does on every occasion is what he is able to do. Nonetheless, this requirement provides fairly objective criteria. The problem of inference criteria for Substage III-A is more difficult, because the Substage III-A subject is supposed to give evidence of formal operational thinking while being unsuccessful on the task or tasks. This is where subjective decisions enter--if a subject is unsuccessful on a task, is the structure absent, or is it present but not yet stable or fully developed?

For the present purposes the problem has to be considered in relation to the idea of intellectual structural limitations, where the inference of lack of structures is taken to imply restrictions on curriculum content and teaching method. In the application of this idea, Type II errors may be seen to be more costly than Type I errors. Sweeping changes to secondary school curricula are being proposed on the basis of an inference that most students lack formal operational structures. For example, according to Lawson and Renner, if such an inference is correct, then "the curriculum will have to

be reshaped to reflect concrete objects, events, and situations--rather than abstract mental images, rules, and roles--in the teaching methods and materials used [1974, p. 558]." To implement major changes in curriculum and method when the inference of absence of structure may be incorrect would be imprudent. Therefore for the purposes of testing the idea of intellectual structural limitations, Type II errors should be avoided in classifying subjects as Substage III-A or Substage III-B. In other words, the conservative course of action in this case is to adopt rigorous criteria of failure.

In the present study four tests were used: two tests to classify subjects according to each of Substages III-A and III-B. In order to classify a subject into a substage on the basis of two tests, two possible criteria could be used: success on both, or success on either.

Suppose that a task T is used to determine whether a given subject has a structure S. Suppose further that a certain type of behavior B on T is considered to be evidence for S. If the subject actually has S but the experimenter judges that he does not display B on T, then a Type II error will be made by the experimenter. If on the other hand the subject is judged to display B but does not have S, then a Type I error will be made.

Such error may be random or systematic. In order to classify a subject into a substage on the basis of two tasks, two possible criteria could be used: success on either, or success on both. The following argument shows that the criterion success on either is subject to more Type I random error but less Type II random error than the criterion success on both. A further argument will be advanced to show that since systematic Type II error can never be removed in principle, the criterion success on either is preferable on theoretical grounds.

First consider random error. Suppose there are probabilities P_I and P_I' that Type I (random) errors are made on the first and second tasks respectively. Suppose there are probabilities P_{II} and P_{II}' that Type II (random) errors are made on the first and second tasks respectively. P_I will be the expected fraction of the subjects who should have passed the first task but failed, and P_{II} will be the expected fraction of the subjects who should have failed the first task but passed. If success on both tasks is adopted as the criterion, the overall probability of Type I errors is $P_I \cdot P_I'$, while the overall probability of Type II errors is $P_{II}+P_{II}'-P_{II} \cdot P_{II}'$. If success on either task is adopted as the criterion, the overall probability of Type I errors is $P_I+P_I'-P_I \cdot P_I'$, while the overall probability of Type II errors is

$P_{II} \cdot P'_{II}$. Therefore if Type II errors are more costly, the conservative course is to adopt success on either task as the criterion, as long as precautions have been taken to reduce systematic Type I error in each of the tests individually.

In high jump and similar athletic contests, the criterion success on any one of three tries is universally adopted, since the question "Can the jumper clear two metres?" is answered affirmatively if he jumps two metres once. The criterion in any one task of those employed here is analogous to the crossbar on the high jump stands.

Now consider systematic error. Although a structure S may be a necessary condition for a subject to display behavior B on task T, S may not be a sufficient condition for B. Subjects may fail on T not because S is absent, but because of such things as misleading or ambiguous instructions, failure to comprehend what is required, lack of motivation. Therefore while presence of B can be taken to imply presence of S, absence of B can never be taken to imply absence of S.

As Braine (1968) pointed out, "It is clear that if one seeks to state an age at which a particular type of response develops, the only age which is not completely arbitrary is the earliest age at which this type of

response can be elicited using the simplest experimental procedures [p. 187]." In the present case, the aim is to determine whether a given subject is capable of a particular type of response. The only criterion which is not completely arbitrary in this case is whether this type of response can be elicited using the simplest experimental procedures. Operationally, the "simplest experimental procedures" become that task among the ones employed which is simplest for the given subject.

In the present study, only subjects who failed both of the pair of tests for a given substage were classified as being below that substage.

Task selection

That part of the criterion problem related to design or selection of tasks amounts to the problem of task validity. The validity of employing volume conservation, controlling variables, and generating combinations as indicators of Substages III-A or III-B will now be discussed, starting with volume conservation.

According to Piaget, volume conservation "is not worked out conceptually before the beginnings of the formal level" because "without a doubt . . . the conservation of volume throughout changes of form presupposes the ability to handle proportions [Inhelder & Piaget, 1958, p. 36]." This latter ability is supposed to be based

on the formal operational four-group structure (Inhelder & Piaget, 1958, pp. 176-181). An advantage of volume conservation is that a fairly straightforward and objective test may be used to detect the acquisition.

Controlling variables seems to provide both a sufficient indication of the beginnings of Substage III-A and the attainment of Substage III-B. In examining the mechanism by which subjects move from concrete operations to formal operations, Piaget has pointed to realization of the need to control variables as a turning point (Inhelder & Piaget, 1958, pp. 282-283). This realization seems to be characteristically present in formal operational subjects, and characteristically absent in concrete operational subjects. Apparently, when two independent variables are involved in a situation, not only do concrete operational subjects fail to control one of the independent variables in order to test the effect of the other on the dependent variable, "but the idea of doing so does not even occur to them [p. 285]." Moreover, Piaget claims that the use of the method of proof based on controlling variables is by itself alone a sure indicator of the combinatorial system, one of the characteristic structures on Stage III (Inhelder & Piaget, 1958, p. 244). This implies that recognition of the need to control variables is a sufficient condition for Substage III-A.

Furthermore, the scheme of "all other things being equal" is supposed to manifest itself differently in Substage III-A and Substage III-B. "As a general rule and in its authentic form, the schema 'all other things being equal' appears only at Substage III-B [Inhelder & Piaget, 1958, p. 43]." At Substage III-A this type of proof "is still only partially understood [p. 12]." For these reasons, manifestation of the idea of controlling variables seems to be an appropriate criterion for the genesis of formal operations, and ability to use the scheme an appropriate criterion for full realization of formal operations.

Another test for Substage III-B is provided by the combinatorial system. This structure enables a subject to generate systematically all of the possible combinations of a set of objects (the set of all possible subsets). The use of the combinatorial system as a test for Substage III-A is somewhat difficult, because at this level the subject is supposed to proceed in some systematic manner to generate combinations, but without success. There is some subjective difficulty in deciding whether a subject is proceeding systematically or is randomly generating the combinations. On the other hand, the criterion for Substage III-B may be straightforward and objective. The subject may be given a set of elements and asked to combine them in all possible ways.

If the subject succeeds in generating all of the possible combinations given sufficient time, then that subject may be classified as Substage III-B.

Hence two tests of controlling variables, a volume conservation task, and a combinatorial system task were used in the study. The actual forms of these tests will now be described. In subsequent sections, the procedures for administration of the tests, the sampling procedure, and the instructional sequences will be described.

VOLUME CONSERVATION TEST

The volume conservation test to be denoted VC, is a group administered version of a procedure previously designed by the experimenter to minimize systematic Type I and Type II errors. An extended rationale for the form of the test is given elsewhere (Hobbs, 1975). In order to be classified as Substage III-A on this task subjects have to predict correctly that when a ball of clay is transformed in shape (rolled into a "sausage," broken into pieces, or flattened into a "pancake"), and then immersed in a glass of water, the water level will be the same as for an initially identical untransformed ball placed in an identical glass of water. In order to minimize Type I error due to subjects correctly predicting the water

level on the basis of conservation of weight, the testing procedure employs a training sequence designed to illustrate to subjects that when a body is immersed in water, the occupied volume rather than the weight determines the displacement of the water level, and a test to determine which subjects do not understand this. The training procedure also teaches subjects how to respond on the answer sheet.

In order to minimize Type II error due to the use of relational terms such as "more" "same" or "less" the question subjects have to answer is "Where will the water come to?" In addition, subjects respond nonverbally by making a line on a diagrammatic representation of the glass to indicate their predicted level. Instructions for the tester and the response sheet are reproduced in Appendix A (pp. 113-125)

CONTROLLING VARIABLES TESTS

Two tests of controlling variables were used. The first, to be denoted "CV-A," was used to classify subjects into Substage III-A. The second, to be denoted "CV-B," was used to classify subjects into Substage III-B.

CV-A was designed to determine whether a given subject recognizes the need to control variables. Subjects were shown two metal rods: a brass rod 51.4 cm

long, 5 mm in diameter; and an aluminum rod 45.5 cm long, 12 mm in diameter. Each subject was supplied with a printed sheet describing an experiment in which a student attempts to bend the two rods. The test form is reproduced in Appendix A (p. 121). By means of a diagram subjects were shown that the brass rod bends far more than the aluminum rod. Subjects were asked whether this proves that rods made of brass bend more easily than rods made of aluminum, and why or why not. A further question required subjects to suggest two ways to improve the experiment. In order to be classified as Substage III-A, a subject had to give evidence of seeing the need to control at least one of length, diameter, or bending force applied.

CV-B was designed to test whether a subject could devise a procedure to test the effect of a single variable while holding a number of other variables constant. In the Inhelder-Piaget (1958, Ch. 3) experiments, subjects were presented with a series of metal rods of varying diameter, cross-sectional shape, and material. The rods were clamped at one end so as to project horizontally. The lengths projecting from the clamp could be varied, and different weights could be attached at the ends of the rods. Subjects were required to test the effects of any one of the variables on the flexibility of a rod, measured by the amount of bending which occurs when weights

are hung on the end.

CV-B is a variation of this test designed to be administered to groups of subjects and objectively scored. Subjects were supplied with a set of seven wooden planks varying according to type of wood (pine or redwood), length (20 cm or 30 cm), width (2.3 cm or 4.5 cm), and thickness (7 mm or 14 mm). The rods were labelled A to G. The redwood planks were clearly different in color from the pine planks, and in addition were clearly labelled as redwood or pine.

Subjects were given a written test to accompany the set of planks. This test is reproduced in Appendix A, together with the instructions for the tester (pp. 113-125). With the aid of a diagram, subjects were shown a method of testing for "stiffness" of a plank, by placing the plank on two blocks and setting a weight on the middle. In a series of three written questions, subjects were then asked how they would test whether the stiffness was affected by length of plank, thickness of plank, or type of wood respectively, and to name the planks they would use. In each case, only one pair of planks would give a valid test properly controlling variables.

A subject could thus select none, one, two, or three pairs correctly. The set of planks is so constructed that three are pine, four redwood; three are

long, four short; three are wide, four narrow; and three are thin, four thick. There are consequently 12 possible ways to pick a pair of planks which differ on one given variable. A subject has three chances to select a correct pair (three questions), so that the probability of selecting exactly one pair correctly by chance is

$$3 \times \frac{1}{12} = \frac{1}{4} .$$

There are three different sets of three responses containing exactly two correct pairs. Therefore the probability of selecting exactly two pairs correctly by chance is

$$\frac{1}{12} \times \frac{1}{12} \times 3 = \frac{1}{48}.$$

The probability of selecting exactly three pairs correctly by chance is

$$\frac{1}{12} \times \frac{1}{12} \times \frac{1}{12} = \frac{1}{1728} .$$

If the criterion for Substage III-B is set at two pairs chosen correctly, then the proportion of all subjects classified incorrectly as Substage III-B would have an expected value of only 0.02. On the other hand, if there is a finite probability p that a given Substage III-B subject might make a mistake in selecting or writing down a pair, the probability that such a subject

will be incorrectly classified as Substage III-A is $3p$ with a criterion of three pairs correct, and approximately $3p^2$ with a criterion of two pairs correct.[1]

Thus using a criterion of three pairs correct, Type I errors are unlikely, but the probability of Type II errors may be quite high. For a criterion of two pairs correct, Type I errors remain virtually negligible, but the probability of Type II errors is reduced virtually by a factor of itself. On the basis of this analysis, the criterion used for classification into Substage III-B on CV-B was set at two pairs selected correctly.

COMBINATORIAL SYSTEM TEST

In the combinatorial system test used by Inhelder and Piaget (1958, Ch. 7), subjects were presented with four different chemical solutions in flasks, and another in a dropper bottle. Subjects were shown that when some of the chemicals were mixed in a certain way, a yellow colour was formed on addition of the liquid from the dropper bottle. They were then required to reproduce the yellow colour by mixing the chemicals. Subjects who proceeded systematically to generate combinations of the

---

[1]Actually the probability is $3p^2\ (1 - \frac{2}{3}p)$.

four solutions were classified as Substage III-A. The more systematic subjects were classified as Substage III-B, the difference being "only one of degree [p. 120]."

The difficulty with using the combinatorial system test to determine Substage III-A lies in the subjectivity of the judgment as to whether a subject is proceeding systematically to generate the combinations or is randomly combining the solutions. The test used in the present study to be denoted CS was designed on the assumption that given a time limit of 15 minutes, only Substage III-B subjects can generate all possible subsets of a given set of four elements. In pilot administrations of CS, after 15 minutes subjects were found either to have finished, or to be duplicating combinations. When a shorter time was used (10 minutes), some subjects complained that they needed more time.

Subjects were given a printed question sheet and answer sheet. These sheets are reproduced in Appendix A (pp.113-125). The problem which subjects had to consider was the number of different ways a transistor radio can be bought given the possibility of up to four "extras" (earphone, carrying strap, leather case, or extra battery). Subjects were required to write or draw "as many different choices they could think of" on the answer sheet. A time limit of 15 minutes was imposed.

There are 15 possible ways of combining the extras,as well as the possibility of no extras. Subjects who generated all 16 possibilities, and subjects who generated 15 not including the possibility of no extras, were classified as Substage III-B. Pilot research involving this test indicated that most subjects who could systematically generate all combinations seemed to feel that the null subset (radio alone) "went without saying."

GENERAL FEATURES OF THE TESTS

The four tests have the following features in common:

(a) Each test is designed to classify subjects according to one substage only, either Substage III-A or III-B.

(b) The criterion for success in each case is objective.

(c) The tests may be administered to groups of subjects.

(d) Subjects record their own responses with pen or pencil on the answer sheets.

(e) Each test takes from 12 to 20 minutes to administer.

Having each test designed to determine only one stage level allows the adoption of objective criteria. These objective criteria are as follows:

In the volume conservation test, the subject had to draw lines representing predicted water levels in four instances. If a subject drew all four lines correctly he was classified as Substage III-A. A line was considered to have been drawn correctly if any part of the line was within 0.5 mm of the correct level. This was determined by drawing a line 1 mm thick at the correct level.

In CV-A, the subject had to indicate whether the test described proves that rods made of brass bend more easily than rods made of aluminum, to say why or why not, and to suggest two ways of making it a better experiment to test whether rods made of brass bend more easily than rods made of aluminum. A subject was classified as Substage III-A if either or both of the following:

(a) he answered "no" to the first question,

(b) he suggested in answer to the third question that at least one of length, diameter, or bending force be equalized for the two rods.

In CV-B, the subject was asked how to test for the effects of length, thickness, and type of wood respectively on the stiffness of a plank, and to indicate

which planks to use. A subject was classified as Substage III-B if in his answers he selected a correct pair of planks at least twice.

In the combinatorial system test, the subject was required to indicate in a 15-minute time limit all the different ways a radio could be bought given four possible "extras." A subject was classified as Substage III-B if he generated all 16 possible ways, or 15 ways not including the radio alone.

ADMINISTRATION OF THE TESTS

Tests were administered to subjects in groups of seven to fourteen. The volume conservation test was administered verbally according to predefined protocols together with demonstrations. The other tests were in written form, but the tester read the instructions and questions to subjects in each case. Instructions for testers were prepared in advance and these are reproduced in Appendix A (pp. 113-125).

THE SAMPLES

A stratified random sample of grade 7 students and a stratified random sample of grade 9 students were selected from the populations of grade 7 student and grade 9 students in three rural high schools in the

Annapolis Valley of Nova Scotia. The grade 7 population consisted of 529 potential subjects. The grade 9 population consisted of 455 potential subjects. Fruit-growing, small farming, and tourism are the major industries in this predominantly white area. The schools in the region adopt a practice of streaming classes according to IQ. The less able students in the lower streams pursue "modified" or "adjusted" programs in smaller classes. In order for the samples to be representative of the populations with respect to ability, sex, and educational experience, subjects were selected randomly from a number of strata. Most strata consisted of all subjects of the same sex in the same class. In order to preserve proportionality, small classes were combined to form single strata in two cases, and males and females in a small class were combined in one case.

The number of students of each sex in each of the strata, and the number selected from each stratum are shown in Tables 1 and 2. The grade 7 sample consisted of 57 subjects, 30 male and 27 female. This represents a sample to population ratio of 0.12. The grade 9 sample consisted of 54 subjects, 25 male and 29 female. This also represents a sample to population ratio of 0.12. In the random selection, more than the required number

TABLE 1

Composition of Strata and Sample

Grade 7

School	Section	Population Males	Population Females	Sample Males	Sample Females
Central Kings	1	22	16	2	2
	2	17	18	2	2
	3	16	16	2	2
	4	17	18	2	2
	5	17	20	2	2
	6	10	15	2	2
	7	10	4		
Hants West	A	15	20	2	2
	B	21	18	2	2
	C	20	17	2	2
	D	21	16	2	2
	E	13	5	2	1
	F	14	2		
Cornwallis	A	14	18	2	2
	B	16	13	2	1
	C	12	11	1	1
	D	16	10	2	1
	E	11	10	1	1
Total		282	247	30	27

TABLE 2

Composition of Strata and Sample

Grade 9

School	Section	Population		Sample	
		Males	Females	Males	Females
Central Kings	16	16	16	2	2
	17	17	15	2	2
	18	14	14	2	2
	19	15	15	2	2
	20	17	10	2	1
	21	6	5	1	-
Hants West	A	15	20	2	2
	B	10	20	1	2
	C	17	12	2	2
	D	13	17	2	2
	E	12	9	1	1
Cornwallis	A	15	19	2	2
	B	15	19	2	2
	C	12	22	1	3
	D	11	17	1	2
	E	5	15	-	2
Total		210	245	25	29

of subjects were selected in each stratum, and the potential subjects were listed in order of selection. Subjects who were absent for one of the pretests were replaced by the next available member on the list for that stratum. No replacements were made after the pretests.

Individual IQ scores were not available for the subjects, but a frequency distribution according to Lorge-Thorndike IQ was obtained for most of the grade 7 subjects who took the posttests. The frequency distribution is shown in Table 3. A chi-square goodness of fit test indicates that the frequency distribution is not significantly different from the normal distribution ($X^2 = 4.91$, $df = 5$, $p > 0.30$).

INSTRUCTIONAL SEQUENCES

Two instructional sequences were used: The combinatorial sequence was designed to teach a systematic method of generating the set of all subsets of a set of four elements. The controlling variables sequence was designed to teach (a) the need to control variables, and (b) the skill of selecting test pairs from given arrays. Full details of instructions for the teachers of the sequences are given in Appendix B (pp. 126-137).

TABLE 3

Frequency Distribution of IQ of Grade 7 Sample and Goodness of Fit to Normal Curve

IQ Range	Observed	Expected	Chi-Square
0 - 80	4	4.8	.13
81 - 90	6	8.5	.74
91 - 100	15	13.2	.25
101 - 110	9	13.2	1.34
111 - 120	12	8.5	1.44
121 -	7	4.5	1.01
Total	53	53	4.91

$df = 5$, $p > 0.30$

The combinatorial sequence

The combinatorial sequence consisted of three 20-minute sessions. In the first session, each subject was presented with 20 paper baking cups and an envelope containing 40 cardboard triangles of varying size and shape, 10 blue, 10 red, 10 yellow, and 10 black. Sides of the triangles varied from about 2 cm to 4 cm. Subjects were given the task of placing one, two, three, or four triangles in the cups in as many different ways as possible, so that no two cups had the same contents, and no cup contained more than one triangle of any one colour. In pilot testing the use of cups for the combinations was found to reduce subjects' tendency to consider permutations rather than combinations. As the subjects attempted the task, the instructor supervised subjects' activity, clarifying the rules, and pointing out incorrect subsets or duplicates. The objective of this session was to make sure that subjects understood the problem in preparation for being given a systematic method of solving it.

In the second session, the instructor explained a systematic way to generate the complete set of combinations by ones, twos, threes, and fours. Subjects were then allowed to practise the systematic method, first with the cardboard triangles previously used,

then with kits of four different kinds of objects--10 each of paper clips, paper fasteners, "bingo chips," and pennies. Instead of baking cups, pink sheets ruled into 20 squares were used. Subjects were allowed to keep the pennies. The systematic method consisted of placing the objects in a row, then taking successive pairs as follows: first and second, first and third, first and fourth, second and third, second and fourth, third and fourth. Sets of three were then generated by omitting one of the four in turn, and finally all four were taken together.

In the third session, subjects were given printed sheets giving practice in generating combinations using pencil and paper rather than manipulating real objects. These printed sheets are reproduced in Appendix B (pp. 126-137). In the first part of the practice session, subjects were given a choice of three problems. One required the generation of all possible "outfits" given four different items of clothing. Another required the generation of all possible combinations of ketchup, relish, cheese, or onions on a hamburger. The other required the generation of all possible combinations of whitewall tires, automatic transmission, radio, or car heater in ordering a new automobile.

In the second part of the practice session subjects were required to draw "all the different ways you can have a square" if the square can have any of an inscribed circle, a dot in the centre, a border, or a cross of diagonals. At no time throughout the combinatorial sequence was any mention made of the transistor radio problem.

Controlling variables sequence

Like the combinatorial sequence, the controlling variables sequence consisted of three 20-minute sessions. The first session was designed to teach subjects the need to control variables. The instructor demonstrated two experiments where control of variables was necessary but not carried out, and then used the results to draw false unlikely conclusions.

In the first demonstration, two glasses were shown to subjects, one 9 cm high and 8 cm in diameter at the top, the other 11 cm high and 6 cm in diameter at the top. A labelled container of sugar, a teaspoon, and a tablespoon were placed on the table beside the the glass. The instructor demonstrated that "sugar dissolves quicker in a short glass than in a tall glass," by pouring water into the two glasses to different levels and stirring a large amount of sugar from the tablespoon

in the tall glass and a small amount of sugar from the teaspoon in the short glass. A discussion ensued as subjects objected that it was not a proper proof, and that both equal volumes of water and equal amounts of sugar would be necessary to "do it right."

The second demonstration involved "proving" that red coloured rubber is "stretchier" than blue coloured rubber by hanging a heavy weight on a long thick red rubber band and a light weight on a short thin blue rubber band. A discussion ensued with the conclusion that equal weights, and rubber bands of the same length and thickness, would be needed to perform a proper test.

In the second session of the controlling variables sequence subjects were shown a photographic light meter. Its use was demonstrated using a light bulb. An array of light bulbs of different shapes and wattages, some frosted and some clear, were shown to subjects. The array of light bulbs is illustrated in Fig. 1. Subjects were required to pick out pairs of bulbs to be used to test whether colour (frosted or clear), wattage, or shape affect the light meter reading. The remainder of the second session and the third session were designed to give subjects practice in selecting test pairs from given arrays. In session two, each subject was given a packet containing eight strips of paper; one blue and

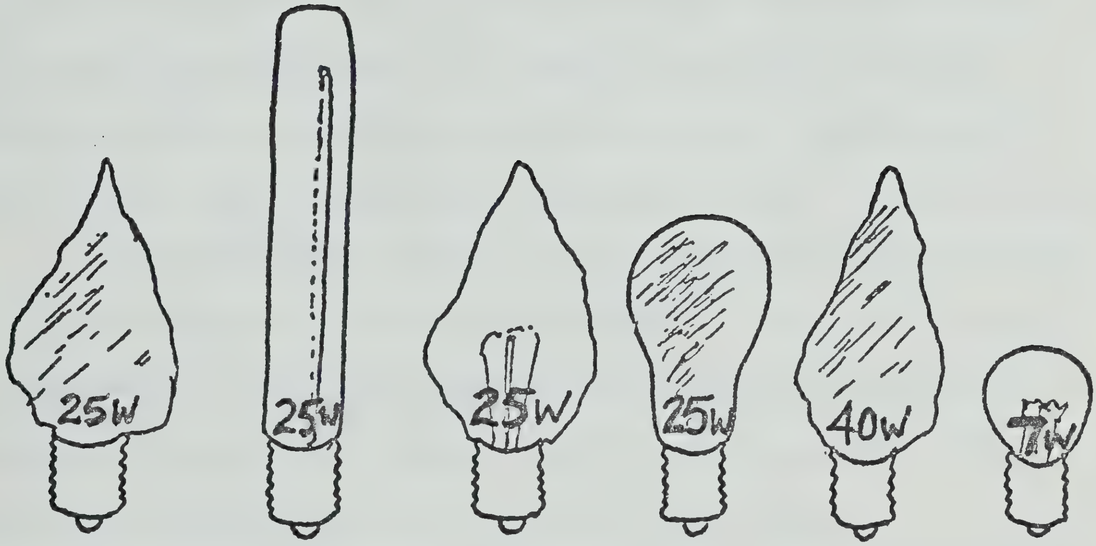


Figure 1

Array of light bulbs used in controlling variables sequence

one pink strip of each of the following four sizes:

13.5 cm x 2.5 cm

13.5 cm x 1.25 cm

8.0 cm x 2.5 cm

8.0 cm x 1.25 cm.

Subjects were asked to select pairs "the same colour, the same length, but different width" and so on. The session concluded with similar practice using (a) rubber tubes of different colour, length, and diameter; (b) metal rods of different material, length, and diameter.

In the third session of the controlling variables sequence subjects were given practice in selecting pairs differing in only one aspect from pictures of arrays of objects varying in a number of different aspects. The pictures and printed questions are reproduced in Appendix B (pp. 126-137). In this case, subjects worked with pen on paper rather than with real objects. Throughout the controlling variables sequence no mention was made of the problem of the brass and aluminum bars nor the problem of the wooden planks (CV-A and CV-B).

PROCEDURE

Tests and instruction sessions were administered to subjects in their respective schools so that any one test or instructional sequence was given to all subjects on the same day but at different times. The schedule

schedule in Table 4 was followed for the grade 7 sample. In each session the subjects were divided into two equal or nearly equal groups. Two experimenters were always present. Each group spent 20 minutes with each experimenter to make a total of 40 minutes for each session. Each experimenter administered the same test or the same instructional session to all subjects in the one day.

Five different experimenters (coded A to E in the schedule in Table 4) took part at various times according to availability. All experimenters were male. All were experienced in teaching secondary school students and in administering tests.

The four tests were administered to the grade 9 sample on April 10, 1975. One experimenter administered the volume conservation test and the CV-B test, the other experimenter administered the CV-A test and the combinatorial system test.

SUMMARY

Samples of 57 grade 7 students and 54 grade 9 students were selected from the populations of grade 7 and grade 9 students in three rural high schools in Nova Scotia. A test of volume conservation, two tests of controlling variables, and a test of generating combinations were administered to both samples. Subjects were

TABLE 4

Testing and Instruction Schedule for Grade 7

Session Number	Date	Experimenters	Test or Instruction
1	February 18	A	Pretest volume conservation
		B	Pretest CV-A
2	February 24	A	Pretest CV-B
		C	Pretest combinatorial system
3	February 27	A	Session No. 1, Controlling variables sequence
		C	Session No. 1, Combinatorial sequence
4	February 28	A	Session No. 2, Controlling variables sequence
		B	Session No. 2, Combinatorial sequence
5	March 5	D	Session No. 3, Controlling variables sequence
		E	Session No. 3, Combinatorial sequence
6	March 7	D	Posttest CV-A and CV-B
		E	Posttest Combinatorial system
7	April 7	A	Retention test CV-A and CV-B
		D	Retention test Combinatorial system

classified as Substage III-A or Substage III-B on the basis of these tests. Each test was designed to reduce the possibility of systematic Type I error and objective criteria were used to score the tests. The grade 7 sample was then subjected to a total of two hours of instruction in controlling variables and generating combinations over a one-week period, and retested for Substage III-B immediately after and one month after the instructional sequence.

CHAPTER IV

RESULTS

Complete results are shown in Table 5 for the grade 7 sample, and in Table 6 for the grade 9 sample. The proportions of subjects who met the criteria established in Chapter III for success on individual tasks and for classification into Substages III-A and III-B will now be considered. Subsequent sections of this chapter will be concerned with changes in subjects' performance on tasks following the instructional sequences and the retention period, and with correlations among the tasks. The chapter concludes with a discussion of the results.

PERFORMANCE ON TESTS

Percentages of subjects who were successful on each task are shown in Table 7. Percentages of subjects who met Substage III-A and Substage III-B criteria are shown in Table 8. Before instruction, 51% of the grade 7 subjects and 80% of the grade 9 subjects passed either or both of VC and CV-A and were classified as Substage III-A. Before instruction, 47% of the grade 7 subjects and 50%

TABLE 5

Results for Grade 7 Sample[1]

Subject Number	Sex	Age (months)	Pretest				Posttest			Retention Test		
			VC	CV-A	CV-B	CS	CV-A	CV-B	CS	CV-A	CV-B	CS
1	M	163	1	1	1	0	1	0	0	1	1	1
2	F	157	0	0	1	0	0	1	0	0	0	0
3	F	150	0	0	0	1	0	0	0	0	0	1
4	F	154	1	1	0	0	1	1	1	1	1	1
5	M	151	0	1	1	0	a	a	a	1	1	0
6	F	159	1	1	0	1	1	1	0	1	1	1
7	M	174	1	0	0	0	1	1	1	a	a	a
8	F	155	1	0	0	0	0	1	1	1	1	1
9	F	155	1	0	0	0	0	1	0	0	1	0
10	M	179	0	0	a	a	a	a	a	a	a	a
11	M	159	1	1	1	0	1	1	1	1	1	1
12	F		a	a	0	0	0	1	0	1	0	0
13	F	164	1	0	1	0	0	0	0	1	0	0
14	M	151	0	0	0	0	0	1	1	a	a	a
15	F	181	0	0	a	a	a	a	a	a	a	a

TABLE 5 (cont'd)

Subject Number	Sex	Age (months)	Pretest				Posttest			Retention Test		
			VC	CV-A	CV-B	CS	CV-A	CV-B	CS	CV-A	CV-B	CS
16	F	153	1	1	0	1	1	1	1	1	0	1
17	M	150	1	1	0	1	1	1	0	1	1	0
18	M	187	0	0	0	0	a	a	a	0	0	0
19	F	148	0	0	0	0	1	0	1	1	1	1
20	M	155	0	0	1	0	a	a	a	0	1	0
21	F	150	0	1	a	a	1	1	0	1	1	0
22	F	155	0	0	0	0	0	0	0	a	a	a
23	M	157	0	1	0	1	1	1	1	a	a	a
24	F	176	a	a	1	0	1	0	0	a	a	a
25	M	169	0	0	0	0	1	1	1	1	1	1
26	F	154	0	1	0	0	1	0	0	1	0	0
27	F	165	0	0	1	0	0	1	0	a	a	a
28	M	157	1	1	1	1	1	1	1	a	a	a
29	F	152	1	0	0	0	0	0	1	1	0	1
30	F	164	0	0	0	0	0	0	0	0	1	0
31	F	156	a	a	1	0	1	1	1	1	1	1

TABLE 5 (cont'd)

Subject Number	Sex	Age (months)	Pretest				Posttest			Retention Test		
			VC	CV-A	CV-B	CS	CV-A	CV-B	CS	CV-A	CV-B	CS
32	M	150	0	0	a	a	a	a	a	a	a	a
33	F	174	0	0	0	0	0	0	0	1	0	0
34	M	171	0	0	1	0	0	1	1	1	1	0
35	M	178	0	0	0	0	0	0	0	0	0	0
36	F	164	0	0	0	0	0	1	0	1	0	0
37	M	153	1	1	1	0	1	0	1	1	1	0
38	F	160	0	0	0	0	1	0	0	a	a	a
39	F	160	1	0	1	0	0	1	0	1	1	0
40	F	168	1	0	0	0	a	a	a	1	0	0
41	M	151	a	a	0	0	1	0	1	1	0	1
42	M	159	0	1	1	0	1	1	1	1	1	1
43	M	181	1	0	0	0	a	a	a	1	1	0
44	M	157	1	1	1	0	1	1	0	1	1	1
45	M	152	1	0	1	0	0	1	1	0	1	1
46	F	165	0	0	0	1	0	1	1	a	a	a
47	M	171	1	0	0	0	a	a	a	1	0	0
48	F	163	0	0	0	0	0	0	0	1	1	0

TABLE 5 (cont'd)

Subject Number	Sex	Age (months)	Pretest				Posttest			Retention Test		
			VC	CV-A	CV-B	CS	CV-A	CV-B	CS	CV-A	CV-B	CS
49	M	160	1	1	1	0	a	a	a	0	0	1
50	M	155	0	0	0	0	1	0	0	1	0	0
51	M	161	0	0	0	0	0	1	1	0	0	1
52	M	177	1	0	0	0	0	0	0	0	0	0
53	F	149	0	1	1	0	1	1	1	1	1	1
54	M	160	0	0	0	1	0	0	1	a	a	a
55	M	169	0	0	0	1	1	0	1	0	0	1
56	M	160	0	1	1	0	1	1	0	1	1	1
57	M	187	a	a	0	0	0	0	0	0	0	0
58	F	149	0	0	0	0	0	1	1	0	0	1
59	M	158	1	1	0	0	0	0	0	1	0	0
60	F	169	0	0	a	a	a	a	a	a	a	a
61	F	170	0	0	0	0	0	0	0	0	0	0
62	M	151	0	0	0	0	1	1	1	1	1	0

[1] 1 = pass, 0 = fail, a = absent. Subjects absent for CV-B and CS pretest were replaced by other subjects and their results not used.

TABLE 6

Results for Grade 9 Sample[1]

Subject Number	Sex	VC	CV-A	CV-B	CS
1	M	2	2	2	1
2	M	2	2	1	1
3	F	2	2	2	2
4	F	2	2	2	2
5	M	1	1	1	1
6	F	2	2	1	1
7	F	2	2	2	2
8	M	1	2	1	1
9	M	2	2	2	2
10	F	1	2	1	1
11	F	1	2	1	1
12	M	1	2	1	2
13	M	2	2	1	1
14	F	2	2	1	2
15	F	1	1	1	1
16	M	1	1	1	1
17	F	1	1	1	1
18	M	2	1	2	1
19	M	2	2	2	1
20	F	2	1	1	1
21	F	2	1	1	2

TABLE 6 (cont'd)

Subject Number	Sex	VC	CV-A	CV-B	CS
22	M	1	2	2	1
23	M	2	2	1	1
24	F	2	2	1	2
25	F	1	2	2	1
26	M	1	2	1	1
27	M	2	2	2	2
28	F	2	2	1	2
29	F	2	2	2	2
30	M	1	1	1	1
31	M	1	2	1	1
32	F	1	1	1	2
33	F	1	1	1	2
34	M	2	2	2	2
35	M	2	2	2	2
36	F	2	1	1	1
37	M	1	2	1	1
38	M	2	2	2	2
39	M	2	2	1	2
40	F	2	2	2	2
41	M	2	2	1	1
42	M	1	2	1	2
43	M	1	2	1	1

TABLE 6 (cont'd)

Subject Number	Sex	VC	CV-A	CV-B	CS
44	F	2	2	2	1
45	F	2	1	2	2
46	M	2	2	1	1
47	F	2	1	1	1
48	F	1	2	1	1
49	F	1	1	1	2
50	M	2	1	1	1
51	F	1	1	1	1
52	F	1	1	1	1
53	F	2	1	1	1
54	F	1	1	1	1

[1] 1 = fail, 2 = pass.

TABLE 7

Per Cent of Subjects Successful on Tasks

Task	Grade 7 Pretest N=57	Grade 7 Posttest N=50	Grade 7 Retention N=47	Grade 9 N=54
Volume Conservation	39			57
CV-A	32	48	64	65
CV-B	33	44	51	31
CS	16	48	45	39

TABLE 8

Frequency of Subjects Classified as Substage III-A and Substage III-B

Test	Met Substage III-A Criteria		Met Substage III-B Criteria		N
	Frequency	Per Cent	Frequency	Per Cent	
Pretest Grade 7	29	51	27	47	57
Posttest Grade 7			34	68	50
Retention Grade 7			32	68	47
Grade 9	43	80	27	50	54

of the grade 9 subjects passed either or both of CV-B and CS and were classified as Substage III-B. (Substages III-A and III-B are not exclusive categories here.)

Comparisons between grade 7 pretest performance and grade 9 performance can readily be made using Tables 7 and 8. A directional chi-square test (Siegel, 1956, pp. 104-111) is convenient to determine whether the proportions meeting criteria in the grade 9 sample are significantly higher than in the grade 7 sample. The grade 9 performance is significantly better than the grade 7 pretest performance on VC ($\chi^2 = 4.36$, $df = 1$, $p < .025$), on CV-A ($\chi^2 = 11.0$, $df = 1$, $p < .0005$), and on CS ($\chi^2 = 6.17$, $df = 1$, $p < .01$). There is no significant difference between the grade 7 pretest preformance and the grade 9 performance on CV-B ($\chi^2 < 10^{-5}$, $df = 1$, $p > .10$).

The proportion of grade 9 subjects meeting the criteria for Substage III-A is significantly higher than the proportion of grade 7 subjects meeting the criteria for Substage III-A ($\chi^2 = 8.92$, $df = 1$, $p < .005$). On the other hand, the proportion meeting the criteria for Substage III-B is not significantly higher for the grade 9 sample than for the grade 7 sample on the pretest ($\chi^2 = .007$, $df = 1$, $p > .10$). More than half of the grade 7 sample and four-fifths of the grade 9 sample met

the Substage III-A criteria. About half of each of the grade 9 sample and the grade 7 sample on the pretest met the Substage III-B criteria.

INSTRUCTION AND RETENTION EFFECTS

For the grade 7 sample, performance on the pretest was compared with performance on the posttest using the McNemar test for significance of changes (Siegel, 1956, pp. 63-67). For each of CV-A, CV-B, CS, and the Substage III-B classification, the following null hypothesis was tested:

H_o: for those subjects who change from pretest to posttest, the probability (P_a) that any subject will change from pass to fail is equal to the probability (P_d) that the subject will change from fail to pass. That is, $P_a = P_d = 1/2$. H_1: $P_d > P_a$.

A directional test was used since subjects were predicted to perform better on the posttest than on the pretest. The significance level was set at five per cent. At this level for a one-tailed test the critical value of chi-square with one degree of freedom is 2.71. A significant chi-square indicates that performance on the posttest was better than performance on the pretest.

Table 9 shows changes in performance on CV-A, CV-B, and CS between the pretest and posttest. The number

TABLE 9

Changes Between Pretest and Posttest[1]

		Posttest Fail	Posttest Pass	χ^2 (McNemar Test)
		CV-A (N = 45)		
Pretest	Pass	1	15	
	Fail	22	7	1.78
		CV-B (N = 50)		
Pretest	Pass	4	12	
	Fail	18	16	6.05*
		CS (N= 50)		
Pretest	Pass	3	6	
	Fail	23	18	9.35*
		Substage III-B Classification (N = 50)		
Pretest	Pass	4	21	
	Fail	12	13	3.71*

*$p < .05$

[1] In this and subsequent tables, only subjects present for both administrations are shown.

of subjects present for both pretest and posttest was 45 for CV-A and 50 for CV-B and CS. Performance improved between pretest and posttest for both CV-B ($\chi^2 = 6.05$, $df = 1$, $p < .01$) and CS ($\chi^2 = 9.35$, $df = 1$, $p < .005$). There was no significant improvement in CV-A ($\chi^2 = 1.75$, $df = 1$, $p > .05$). Table 9 also shows changes in subjects' classification according to Substage III-B criteria between the pretest and posttest. The number of subjects classified as Substage III-B increased from 25 to 34. This improvement is significant ($\chi^2 = 3.71$, $df = 1$, $p < .05$).

Table 10 shows changes in performance between posttest and retention test. Changes are not significant for CV-B ($\chi^2 = 0$, $df = 1$, $p > .10$), nor for CS ($\chi^2 = 0.13$, $df = 1$, $p > .10$). The proportion of subjects classified as Substage III-B did not change significantly between posttest and retention test ($\chi^2 = 0.11$, $df = 1$, $p > .10$). For CV-A, however, more subjects passed the retention test than passed the posttest. The difference is significant using a non-directional test ($\chi^2 = 4.00$, $df = 1$, $p < .05$).

A more direct test was used to determine whether lasting significant gains were achieved in Stage III performance by comparing pretest performance with retention test performance. The results are shown in Table 11. Significant improvement between pretest and

TABLE 10

Changes Between Posttest and Retention Test

		Retention Test Fail	Retention Test Pass	Chi-Square (McNemar Test)
		CV-A (N = 40)		
Posttest	Pass	1	18	
	Fail	13	8	2.33
		CV-B (N = 40)		
Posttest	Pass	6	16	
	Fail	13	5	0.32
		CS (N = 40)		
Posttest	Pass	3	15	
	Fail	17	5	0.70
		Substage III-B Classification (N = 40)		
Posttest	Pass	3	25	
	Fail	9	3	0.11

TABLE 11

Changes Between Pretest and Retention Test

		Retention Test		Chi-Square
		Fail	Pass	(McNemar Test)
		CV-A ($\underline{N}$ = 43)		
Pretest	Pass	2	12	
	Fail	12	17	10.3 *
		CV-B ($\underline{N}$ = 47)		
Pretest	Pass	2	14	
	Fail	20	11	4.9 *
		CS ($\underline{N}$ = 47)		
Pretest	Pass	1	4	
	Fail	25	17	12.5*
		Substage III-B Classification ($\underline{N}$ = 47)		
Pretest	Pass	2	19	
	Fail	13	13	6.66*

*$\underline{p} < .05$

retention test was shown in CV-A (χ^2 = 10.3, *df* = 1, *p* < .005), in CV-B (χ^2 = 4.9 , *df* = 1, *p* < .05), and in CS (χ^2 = 12.5, *df* = 1, *p* < .0005). The proportion of subjects who met the Substage III-B criteria increased significantly from 45% to 68% (χ^2 = 6.66, *df* = 1, *p* < .005).

To summarize, significant gains occurred in performance on CV-B and CS between pretest and posttest. Significant gains occurred in performance on CV-A, CV-B, and CS between pretest and retention test. The proportion of subjects classified as Substage III-B increased significantly from 50% to 68% between pretest and posttest, and from 45% to 68% between pretest and retention test (percentages calculated of subjects who were present for both tests).

One further aspect of the changes in behavior was considered: that of a predicted interaction between age and instruction with Substage III-B as criterion. According to the idea of intellectual structural limitations, those subjects most likely to accelerate to Substage III-B in the period between pretest and posttest would be the older subjects. These subjects would be more likely to be closer to "natural" development of Substage III-B behavior. A two-way analysis of variance with age and instruction as independent variables and performance on tasks as criterion was applied. Performance on tasks was determined by allotting one point to a subject for

success on each of CV-A, CV-B, and CS, making a possible three points. This design has repeated measures on the instruction factor. Subjects were all chosen from one grade level in the study, and since students of lower achievement often repeat years in the educational system which supplied the sample, some precaution was necessary to counteract the possibility that any interaction effect would be masked by older subjects of lower ability. The precautions were first, the removal from the sample of subjects over the age of 169 months (older subjects must have entered school late or else repeated a grade); and second, the removal from the sample of students in the lower ability streams in each school (those streams pursuing modified or adjusted programs). After the exclusion of such subjects, 35 remained who were present for both pretest and posttest. Age was made a fixed factor with two levels: those born before and those born after January 17, 1962. The subject of median age (born January 17, 1962) was excluded, equalizing the cell frequencies. There was no significant interaction between age and instruction ($\underline{F} = 2.00$, $\underline{df} = 1.32$, $\underline{p} > .10$).

CORRELATIONS AMONG TESTS

Tables 12 and 13 are matrices of phi-coefficients of correlations among the tests for the two samples.

TABLE 12

Phi-coefficients Among Tests (Grade 7 Sample)

	Pretest				Posttest			Retention Test		
	VC	CV-A	CV-B	CS	CV-A	CV-B	CS	CV-A	CV-B	CS
Pretest										
VC		.31*	.18	.01	.15	.10	.04	.23	.08	.13
CV-A	.31*		.39*	.22	.67*	.24	.05	.42*	.28*	.30*
CV-B	.18	.39*		-.20	.20	.26	.05	.11	.40*	.17
CS	.01	.22	-.20		.18	.10	.09	-.06	.04	.25
Posttest										
CV-A	.15	.67*	.20	.18		.14	.34*	.52*	.32*	.41*
CV-B	.10	.24	.26	.10	.14		.35*	.19	.45*	.27
CS	.04	.05	.05	.09	.34*	.35*		.14	.23	.61*
Retention Test										
CV-A	.23	.42*	.11	-.06	.52*	.19	.14		.34*	.05
CV-B	.08	.28*	.40*	.04	.32*	.45*	.23	.34*		.17
CS	.13	.30*	.17	.25	.41*	.27	.61*	.05	.17	

*$p < .05$

TABLE 13

Phi-coefficients Among Tests (Grade 9 Sample)

	VC	CV-A	CV-B	CS
VC		.24	.42*	.30*
CV-A	.24		.33*	.19
CV-B	.42*	.33*		.36*
CS	.30*	.19	.36*	

*$p < .05$

For $N = 50$, a phi-coefficient of 0.28 is significantly different from zero ($p < .05$). In the remainder of this section, attention will be confined to those particular cases where positive correlations are to be expected for certain reasons.

A measure of the reliability of the individual tasks and of the Substage III-B criteria can be gained by considering the correlations between posttest and retention test results on the tasks and on the Substage III-B classification. The phi-coefficients for CV-A, CV-B, and CS are respectively 0.59, 0.54, and 0.61. For Substage III-B classification the phi-coefficient is 0.64. The corresponding contingency tables appear in Table 9 (p. 69).

For the classification as Substage III-B, 62.5% met the criteria on both posttest and retention test, 22.5% met the criteria on neither posttest nor retention test, and 15% met the criteria on one test and not the other. In other words, the classification procedure was consistent for 85% of the subjects.

Formal-operational structures are supposed to form an integrated system (Inhelder & Piaget, 1958, Ch. 17), therefore the four formal operational tests would be expected to correlate with each other. In particular, VC and CV-A should correlate with each other. The phi-coefficients between VC and CV-A are 0.33 on the

pretest for the grade 7 sample, and 0.36 for the grade 9 sample. Both are significant ($p < .05$).

CV-A and CV-B should not only correlate, but also be related to each other in hierarchical fashion. Fewer subjects would be expected to fail CV-A and pass CV-B than *vice versa*. Table 14 shows the relationship between CV-A and CV-B performances by the grade 9 sample and on the pretest, posttest, and retention tests by the grade 7 sample. Phi-coefficients are significant except in the posttest for grade 7. To determine if the predicted hierarchical relationship between CV-A and CV-B holds, a test similar to the McNemar test used previously is appropriate. The null hypothesis in this case is:

H_o: For those subjects who pass only one test, the probability that any subject will pass CV-A and fail CV-B (that is P_a) is equal to the probability that any subject will fail CV-A and pass CV-B (that is P_d). That is $P_a = P_d = 1/2$.

H_1: $P_d > P_a$.

Again a directional test is used since the hierarchical order is predicted. The chi-square values are not significant for the grade 7 pretest ($\chi^2 = .071$, $df = 1$, $p > .10$), posttest ($\chi^2 = .41$, $df = 1$, $p > .10$), or retention test ($\chi^2 = 2.4$, $df = 1$, $p > .05$). The

TABLE 14

Relationship Between CV-A and CV-B

		CV-B Fail	CV-B Pass	Phi
		Grade 7 Pretest (N = 52)		
CV-A	Pass	7	10	
	Fail	28	7	.39*
		Grade 7 Posttest (N = 50)		
CV-A	Pass	9	15	
	Fail	13	13	.13
		Grade 7 Retention Test (N = 47)		
CV-A	Pass	12	20	
	Fail	11	4	.39*
		Grade 9 (N = 54)		
CV-A	Pass	20	15	
	Fail	17	2	.35*

*$p < .05$.

chi-square value is significant for the grade 9 test ($\chi^2 = 12.2$, $df = 1$, $p < .0005$).

Significant correlations would also be expected between performance on CV-B and CS, the two Substage III-B tasks. Table 15 shows the relationship between CV-B and CS on the pretest, posttest, and retention tests for grade 7, and for grade 9. Two of the phi-coefficients are significant ($p < .05$): those for the grade 7 posttest (0.33) and for the grade 9 test (0.36).

An hierarchical relationship is expected between Substage III-A classification and Substage III-B classification. Table 16 shows the relationship between the Substage III-A and III-B classifications for grade 7 (pretest) and grade 9. The respective phi-coefficients are not significant ($p > .05$). The predicted hierarchical relationship is borne out in the grade 9 results ($\chi^2 = 12.0$, $df = 1$, $p < .0005$), but not in the grade 7 results ($\chi^2 = 2.0$, $df = 1$, $p > .05$). The proportions of subjects who meet the Substage III-B criteria but not the Substage III-A criteria are 16% of the grade 7 sample, and 6% of the grade 9 sample.

To summarize, the results of the correlational analysis indicate that the tests and classification criteria are fairly reliable. Correlations predicted to be significant are generally so, except that the posttest grade 7 performance on CV-A does not correlate

TABLE 15

Relationship Between CV-B and CS

		CS Fail	CS Pass	Phi
		Grade 7 Pretest (N = 57)		
CV-B	Pass	18	1	
	Fail	30	8	.21
		Grade 7 Posttest (N = 50)		
CV-B	Pass	11	18	
	Fail	15	6	.33*
		Grade 7 Retention Test (N = 47)		
CV-B	Pass	12	13	
	Fail	14	8	.21
		Grade 9 (N = 54)		
CV-B	Pass	6	11	
	Fail	27	10	.36*

*$p < .05$.

TABLE 16

Relationship Between Substage III-A and Substage III-B Classifications

		Substage III-B Fail	Substage III-B Pass	Phi
		Grade 7 ($\underline{N}$ = 52)		
Substage III-A	Pass	11	17	
	Fail	16	8	.26
		Grade 9 ($\underline{N}$ = 54)		
Substage III-A	Pass	19	24	
	Fail	8	3	.23

significantly with the posttest grade 7 performance on CV-B, and that the grade 7 performance on CV-B does not correlate significantly with the grade 7 performance on CS in the pretest or retention test.

Two predictions of hierarchical relationship are not borne out in the grade 7 sample: CV-A was not prerequisite to CV-B in the pretest or the posttest. The Substage III-A classification criteria were not prerequisite to the Substage III-B criteria in the grade 7 sample. On the other hand, hierarchical relationships were borne out quite well in the grade 9 sample. Possible explanations of these particular anomalies will be discussed below, but all of the anomalies seem to be connected with the pretest and posttest grade 7 performance on CV-A.

DISCUSSION

The implications of the results of the study for the idea of intellectual structural limitations are discussed in the remainder of this chapter, together with some of the possible implications of the correlational analysis results for the tests themselves and for the theory of formal operations. The chapter concludes with a note on the generalization of the findings to other populations.

A case has been made that application of the idea of intellectual structural limitations to secondary school rests on the validity of three statements:

(a) The onset of the stage of formal operations (Substage III-A) does not occur for a majority of students until the secondary school years or later.

(b) The full realization of formal operations (Substage III-B) is not attained for a majority of students during the secondary school years.

(c) Progression from Substage III-A to Substage III-B is necessarily slow (of the order of years) and is not subject to appreciable acceleration in the secondary school years.

For structural limitations to be relevant to secondary grades, either statement (a) has to be true, or else both statements (b) and (c) have to be true. Structural limitations are not relevant to secondary grades if statement (a) is false and either of statements (b) or (c) is false.

First consider statement (a). In the survey, 51% of the grade 7 sample and 80% of the grade 9 sample were successful on either VC or CV-A and were classified as Substage III-A. Thus statement (a) is false for the grade 7 and grade 9 samples. For the grade 9 population the statement is false, while for the grade 7 population,

the truth of statement (a) is marginal. Standard errors of the percentages are about 6% for grade 7, 5% for grade 9. A 90% confidence interval around the grade 9 percentage does not contain 50% (80%±8%). A 90% confidence interval around the grade 7 percentage contains 50% (51%±10%). According to the pretest, at least 40% of the grade 7 population had entered the period of formal operations while at most, 30% of the grade 9 population had not entered the period of formal operations. These last two facts (made at the 90% level of confidence), together with the fact that on the retention test 64% (S.E. = 5%) of the grade 7 sample was successful on CV-A, can be taken to indicate that statement (a) is very doubtful.

Statement (b) is that a majority of secondary students are not in Substage III-B. The truth of statement (b) is marginal for both the grade 7 and grade 9 populations. Fifty per cent of the grade 9 sample and 47% of the grade 7 sample were classified as Substage III-B. The standard error of these percentages is about 6%.

Assuming that statement (a) is false and allowing that statement (b) is doubtful but possibly true, the applicability of the idea of intellectual structural limitations to secondary students rests on the truth of statement (c): that appreciable acceleration to Substage III-B is not possible. In the training experiment,

Substage III-B behavior was fairly easily induced in a significant proportion of the grade 7 sample. A majority of the grade 7 students (68%) was classified as Substage III-B both on the posttest and the retention test. This represents a significant increase, from below 50% on the pretest, following two hours of instruction. Of the subjects who took both pretest and retention test, half of those who were classified as being below Substage III-B on the pretest attained Substage III-B classification on the retention test. About one-quarter of the grade 7 sample did not display Substage III-B behavior on any of the test administrations. A reasonable assumption is that alternate and perhaps longer sequences of instruction might be successful in inducing Substage III-B behavior in a larger proportion of grade 7 subjects, or that the same sequence of instruction might be more successful with grade 9 students. Therefore a reasonable conclusion is that if the idea of intellectual structural limitations is intrinsically valid, it can apply to no more than about one-quarter of the grade 7 population, and probably to lesser fractions of corresponding populations in higher grades.

Consider now the results of the correlational analysis. As a general rule, these results are consistent with the structural integrity of formal operations tasks. Significant positive correlations were found

among most tests, although none of the phi coefficients are of any large magnitude. Two anomalies are:

(a) the absence of an hierarchical relationship between CV-A and CV-B, which seemed to carry over into a lack of hierarchical relationship between the Substage III-A and Substage III-B classifications for the grade 7 sample; and

(b) the low correlation between CV-B and CS, the two Substage III-B tasks.

For the grade 7 sample, the CV-A test was anomalous throughout. Although the posttest results on CV-A were not significantly different from the pretest results, the retention test results were significantly better than both. In the posttest, 26% of the subjects failed CV-A but passed CV-B. One interpretation is that the validity of either or both tests is questionable, since CV-A is supposed to determine if subjects see the need to control variables, while CV-B is supposed to determine if subjects are able to control variables. Another interpretation is that there may be no necessary difference between seeing the need to control variables and being able to do so. This interpretation is not borne out, however, by the fact that a strong hierarchical relationship shows up in the grade 9 sample. Possibly the questions asked in CV-A

are too vague and ambiguous, and the level of verbal comprehension demanded is too high for grade 7 subjects. The requirement of verbal responses in this test probably increases the probability of Type II error (Brainerd, 1974). If so, the proportion of Substage III-A subjects in the grade 7 sample was probably underestimated.

The low correlation between CV-B and CS raises theoretical problems. If formal operations are structured in an integrated system, then the results of all formal operational tasks should correlate. For the grade 7 posttest and the grade 9 test the correlations are 0.33 and 0.36. For the grade 7 pretest and retention test, the correlations are both 0.21. The maximum possible values of phi corresponding to the marginal totals in the two cases where significant correlations occur are 0.82 and 0.85, so that the phi-coefficients obtained can be considered to indicate moderate correlation. In the case of the grade 7 pretest, the difficulty of CS probably explains the low coefficient: only nine subjects passed CS and the maximum value of phi for the corresponding marginal totals is only 0.38. The low correlation in the grade 7 retention test remains anomalous. If the four phi-coefficients corresponding to the four tests were averaged, the mean value would be 0.38. Bearing

in mind that the phi-coefficient between posttest and retention test administrations of CV-B and CS were 0.54 and 0.61 respectively, a moderate correlation could be considered to exist between CV-A and CS.

Bredderman (1973) found the results of tests of controlling variables and generating combinations to have a Pearson *r* of 0.49. Bredderman's results include correlation with mental age, and if this is partialled out, the results of the two tests have a partial *r* of 0.31. In a correlational study, Bart (1971) found Pearson coefficients among four Piagetian tests (GLT-3, -8, -11, -13) to range from 0.45 to 0.73. The theory of an integrated structure for formal operations suggests that high correlation coefficients among formal operational tasks should be the rule. The matter warrants further investigation.

Finally, with reference to generalization to other populations, there is no reason to believe that the populations studied contained unusually large proportions of especially bright students, or students who had been exposed to especially rich learning experiences. Therefore the findings of this study can fairly safely be generalized to other populations except where unusual cultural deprivation is found, or where the ability distribution is negatively skewed. The finding that the idea of

intellectual structural limitations does not apply to the grade 7 population may be extended safely to higher grades. There is no reason to suspect that the idea of intellectual structural limitations has any more applicability to normal urban secondary school populations than it has to the population studied.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The study represents an empirical test of the applicability of the idea of intellectual structural limitations to secondary school students, that is that certain kinds of learning are not possible for students in secondary school grades because they lack formal-operational structures and cannot develop them readily.

The research was in two parts, a survey and an experiment. In the survey, samples of grade 7 and grade 9 students from three rural high schools were tested using a volume conservation task, two controlling variables tasks, and a combinatorial system task. In the experiment, the grade 7 sample was then administered two one-hour instructional sequences spaced over a one-week period. One sequence consisted of instruction in controlling variables, and the other consisted of instruction in generating combinations. At the end of the instructional sequence and again one month later, the controlling variables and combinatorial system tasks were readministered.

On the basis of performance on the tests, and in accordance with theoretically derived criteria, 51% of the grade 7 sample and 80% of the grade 9 sample were

classified as being at least in Substage III-A of the period of formal operations; while 47% of the grade 7 sample and 50% of the grade 9 sample were classified as being in Substage III-B of the period of formal operations. After the instructional sequences, 68% of the grade 7 sample was classified as Substage III-B on both the posttest and retention test. The increases in the proportion of Substage III-B subjects from the pretest to the posttest and from the pretest to the retention test are significant at the five per cent level. These results were interpreted to indicate that the idea of intellectual structural limitations does not apply to secondary school grades.

A correlational analysis of the results indicated that the tests are fairly reliable, and that as a general rule, they intercorrelate according to predictions of the theory of formal operations. Exceptions were (a) that the two controlling variables tests failed to display consistently a predicted hierarchical relationship, and (b) that one of the controlling variables tasks and the combinatorial system task, both designed to classify subjects as Substage III-B, did not correlate with each other above a moderate level.

CONCLUSIONS

The results indicate that if the idea of intellectual structural limitations applies to secondary school grades at all, it can apply to no more than about 25% of grade 7 students, and to lower proportions of students in higher grades. Of 34 grade 7 subjects who had not displayed Substage III-B behavior on the controlling variables task in the pretest, 16 displayed such behavior after a total of only one hour's instruction in controlling variables. Of 41 grade 7 subjects who had not displayed Substage III-B behavior on the combinatorial system task in the pretest, 18 displayed such behavior after a total of only one hour's instruction in generating combinations. A reasonable assumption would be that more instruction would be successful with more of the subjects.

If the idea of intellectual structural limitations were to be applied in the selection of learning experiences for the grade 7 sample studied, then in their schooling no attempt would be made to have the abilities to control variables or generate combinations used, applied, or practised by those subjects whose pretest behavior indicates that they do possess these abilities, and no attempt would be made to develop these abilities in those subjects whose pretest and posttest behavior indicate that they can readily acquire the abilities.

There are two possible interpretations of the results alternative to the interpretation given above. The first alternative is that the gains between pretest and retention test are a function of the testing procedure and independent of the development of any abilities. The second alternative is that no structural changes were induced in the instructional sequence but that subjects learned behavior specific to the tests. These alternative interpretations will now be discussed.

In the first alternative, the hypothesis is that in the experiment, subjects learned nothing more than how to do the tests. According to this interpretation, some subjects may have understood the need to control variables, or been able to control variables, or been able to generate combinations, but yet failed the pretest because they did not understand what was required in the tests, or because they were anxious in the test situation, or because they were distracted by the strangeness of the unfamiliar requirements. As these subjects became more familiar with the tests on successive administrations, they performed up to their abilities and went from fail to pass. This interpretation would mean that both statements (a) and (b) are false: that a majority of subjects in secondary school has attained both Substage III-A and Substage III-B as indicated by the results

of the grade 7 retention test. The interpretation further implies that a single administration of a formal-operational task to a group of subjects would always result in underestimation of the proportion of formal-operational subjects.

With the second alternative, the hypothesis is that the Substage III-B performance on the posttest and retention test is specific learned behavior and not the result of structural change in subjects. This would mean that in the same theory, the same behavior allows of two contradictory interpretations. In one case (the pretest) the behavior (B) infers a structure (S), while in another case (the posttest), B infers not S but specific acquired habits. This would mean that except in an entirely novel situation, S is not a necessary condition for B. To make matters worse, S can never be a sufficient condition for B, because a subject cannot be successful on a task (by emitting B) unless he understands what he is being asked to do, is attending, and is willing to cooperate. If S is neither a necessary nor a sufficient condition for a given B, then structure must be a very mysterious entity and the results of Piagetian tasks must be very hazardous bases for important decisions about the kinds of learning experiences to which students should be exposed.

The fact is that at the end of the experiment in the present study, two out of three of the grade 7 subjects

could control variables in testing the stiffness of planks, or generate combinations of accessories to a radio, or both. A total of seven classroom periods--less than a day's school--was taken up in instruction and testing to reach this point. To say that such abilities (being formal-operational) cannot be learned by secondary school students (because they are concrete-operational) is false.

RECOMMENDATIONS FOR FURTHER RESEARCH

The relationship between CV-A and CV-B warrants further investigation: The question is whether there is any necessary time-consuming developmental process between recognition of the need to control variables and being able to control variables in a given situation. In order to answer this question, tests for the recognition of the need to control variables would have to be developed which are less verbal. In order to control for the effect of differing content, CV-A and CV-B items could be paired on the basis of similar content. For example, a Substage III-A controlling variables task involving planks of wood could be paired with a Substage III-B task involving planks of wood. The reliability of each measure could be increased by having a number of items, but this would lead to a problem in keeping the length of the tests down.

The forms of the questions would have to be abbreviated from those used in the present study.

Similarly the relationship between CV-B and CS merits further attention, and controlling for the effect of content would seem to be necessary in this case also. If content is controlled for, then the theory of formal operations would predict that the two measures ought to correlate to a level at least comparable with their reliabilities.

Close attention could well be paid to that minority of subjects who continued to fail to meet the Substage III-A and Substage III-B criteria after the instructional sequences. These subjects could be screened out using the procedures of the present study. An appropriate initial approach to these subjects might consist of clinical interviews or longtitudinal case studies.

A careful attempt to detect the predicted interaction between age and instruction, which did not show up in the experiment, could well be made. Piaget's notion of "natural" development definitely implies that such an interaction should occur. If, as the results of the present study suggest, development in the period of formal operations can be influenced by specific learning experiences, then an interaction between educational

environment and instruction might be more likely. Students from richer educational environments might be less likely to show gains as the result of instruction than students from poorer environments. The design of an experiment oriented to this problem might be quite difficult, but it would represent a clear test between predictions from alternate hypotheses.

Finally, the results of the present study should not be taken to indicate that most secondary students are "fully formal operational" in the sense of being able to solve correctly the problems given in The Growth of Logical Thinking. The facts that the grade 9 sample performed better on the tasks than the grade 8 sample, and that significant training effects were found, indicate that students' thinking ability is developing in the secondary school years. Therefore the final recommendation is for a study of the effects on subjects of varying age and ability levels of long-term instruction in formal-opertional schemes and abilities, such as proportionality, the combinatorial system, controlling variables, and formal logic. Such a study might well result in the finding that contrary to the idea of intellectual structural limitations, the more exposure to and practice in formal thinking, the better.

BIBLIOGRAPHY

Aebli, H. Didactique psychologique. Application a la didactique de la psychologie de Jean Piaget. Neuchatel: Delachaux et Niestlé, 1966.

Almy, M. Young children's thinking. New York: Teachers College Press, 1966.

Anderson, R. C. Can first graders learn an advanced problem-solving skill? Journal of Educational Psychology, 1965, 56, 283-294.

Ausubel, D. P. The psychology of meaningful verbal learning. New York: Grune & Stratton, Inc., 1963.

Ausubel, D. P., & Ausubel, P. Cognitive development in adolescence. Review of Educational Research, 1966, 36, 403-413.

Ausubel, D. P., & Robinson, F. G. School learning an introduction to educational psychology. New York: Holt, Rinehart & Winston Inc., 1969.

Ball, D. W., & Sayre, S. A. Piaget and the secondary science teacher. School Science and Mathematics. 1974, 74, 331-336.

Bart, W. M. The factor structure of formal operations. British Journal of Educational Psychology, 1971, 41, 70-77.

Beard, R. M. Children's reasoning: a repetition of some of Piaget's and Inhelder's experiments by members of the A.T.M. Mathematics Teaching, 1962, 21, 165-228.

Beard, R. M. The order of concept development studies in two fields (II). Educational Review, 1963, 15, 228-237.

Beilin, H. The training and acquisition of logical operations. In Rosskopf, M. F., Steffe, L. P., & Taback, S. (Eds.). Piagetian cognitive-development research and mathematical education. Washington: National Council of Teachers of Mathematics Inc., 1971.

Berzonsky, M. Interdependence of Inhelder's and Piaget's model of logical thinking. Developmental Psychology, 1968, 4, 469-76.

Braine, M. D. S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. In Sigel, I. E., & Hooper, F. H. (Eds.). Logical thinking in children, New York: Holt, Rinehart, & Winston, 1968.

Brainerd, C. J. Continuity and discontinuity hypotheses in studies of conservation. Developmental Psychology, 1970, 3, 225-228.

Brainerd, C. J. The development of the proportionality scheme in children and adolescents. Developmental Psychology, 1971, 5, 469-476.

Brainerd, C. J. Judgments and explanations as criteria for the presence of cognitive structures. Psychological Bulletin, 1973, 79, 172-179.

Brainerd, C. J. Postmortem on judgments, explanations, and Piagetian cognitive structures. Psychological Bulletin, 1974, 81, 70-71.

Brainerd, C. J., & Allen, T. W. Training and generalization of density conservation: effects of feedback and consecutive similar stimuli. Child Development, 1971, 42, 693-704.

Bredderman, T. A. Elementary school science experience and the ability to combine and control variables. Unpublished paper, State University of New York at Stony Brook, 1973a.

Bredderman, T. A. The effects of training on the development of the ability to control variables. Journal of Research in Science Teaching, 1973b, 10, 189-200.

Carreras, R. Épistémologie génétique et enseignement des sciences. In Bresson, F., & de Montmollin, M. (Eds.). Psychologie et épistémologie génétiques thèmes Piagétiens. Paris: Dunod, 1966.

Case, D., & Collinson, J. M. The development of formal thinking in verbal comprehension. British Journal of Educational Psychology, 1962, 32, 103-111.

Case, R. Learning and development. A neo-Piagetian interpretation. Human Development, 1972, 15, 339-358.

Case, R., & Fry, C. Evaluation of an attempt to teach scientific inquiry and criticism in a working class high school. Journal of Research in Science Teaching, 1973, 10, 135-142.

Dale, L. G. The growth of systematic thinking: replication and analysis of Piaget's first chemical experiment. Australian Journal of Psychology, 1970, 22, 277-286.

Dyrli, O. E. D. An investigation into the development of combinatorial mechanisms characteristic of formal reasoning, through experimental problem situations with sixth grade students. (Doctoral dissertation, Indiana University). Ann Arbor, Mich.: University Microfilms, 1967, No. 68-4817.

Egan, D. E., & Grenno, J. G. Acquiring cognitive structure by discovery and rule learning. Journal of Educational Psychology, 1973, 64, 85-97.

Elkind, D. The development of quantitative thinking: a systematic replication of Piaget's studies. Journal of Genetic Psychology, 1961a, 98, 37-46.

Elkind, D. Children's discovery of the conservation of mass, weight, and volume: Piaget replication study II. Journal of Genetic Psychology, 1961b, 98, 219-227.

Elkind, D. Quantity concepts in junior and senior high school students. Child Development, 1961c, 32, 551-560.

Elkind, D. Quantity concepts in college students. Journal of Social Psychology, 1962, 57, 459-465.

Elkind, D. Piaget's conservation problems. Child Development, 1967, 38, 15-27.

Elkind, D., Barocas, R., & Rosenthal, B. Combinatorial thinking in adolescents from graded and ungraded classrooms. Perceptual and Motor Skills, 1968, 27, 1015-1018.

Elkind, D., & Flavell, J. H. (Eds.). Studies in cognitive development. Essays in honor of Jean Piaget. New York: Oxford University Press, 1969.

Ennis, R. H. et al. Conditional logic and children. Ithaca, N. Y.: Cornell Critical Thinking Project, 1969.

Farrell, M. A. The formal stage: a review of the research. Journal of Research and Development in Education, 1969, 3(1), 111-118.

Ferguson, G. A. Statistical Analysis in Psychology and Education. (2nd ed.) New York: McGraw-Hill, 1966.

Fishbein, E., Pampau, I., & Minzat, I. Comparison of ratios and the chance concept in children. Child Development, 1970a, 41, 377-389.

Fishbein, E., Pampu, I., & Minzat, I. Effects of age and instruction on combinatory ability in children. British Journal of Educational Psychology, 1970b, 40, 261-270.

Flavell, J. H. The developmental psychology of Jean Piaget. New Jersey: D. Van Nostrand, 1963.

Friot, F. E. The relationship between an inquiry teaching approach and intellectual development. (Doctoral dissertation, University of Oklahoma). Ann Arbor, Mich.: University Microfilms, 1970. No. 71-12,569.

Goodnow, J. J. A test of milieu effects with some of Piaget's tasks. Psychological monographs, 1962, 76(36, whole No. 555).

Goodnow, J. J., & Bethon, G. Piaget's tasks: the effects of schooling and intelligence. Child Development, 1966, 37, 573-582.

Graves, A. J. Attainment of conservation of mass, weight, and volume in minimally educated adults. Developmental Psychology, 1972, 7, 223.

Green, D. R., Ford, M. P., & Flamer, G. B. Measurement and Piaget. New York: McGraw-Hill, 1971.

Griffiths, J. A., Shantz, C. A., & Sigel, I. E. A. A methodological problem in conservation studies: the use of relational terms. Child Development, 1967, 38, 841-848.

Gyr, J. W., & Fleischer, C. Computer-assisted studies of Piaget's three-stage model of cognitive development. Psychological Reports, 1967, 20, 165-166.

Hale, J. P. An investigation of two formal operational schemata in adolescents enrolled in the I.S.C.S. classrooms of three selected teachers. Paper presented at the 46th annual meeting of the Nationsl Association for Research in Science Teaching, March, 1973.

Hamlyn, D. W. Epistemology and conceptual development. In Mischel, T. (Ed.), Cognitive development and epistemology. New York: Academic Press Inc., 1971.

Henle, M. On the relation between logic and thinking. Psychological Review, 1962, 69, 366-378.

Hermeier, R. K. Age and intelligence as related to the concepts of mass and volume conservation. (Doctoral dissertation, University of Iowa.) Ann Arbor, Mich.: University Microfilms, 1967. No. 67-16,804.

Higgins-Trenk, A., & Gaite, A. J. H. Elusiveness of formal operational thought in adolescents. Proceedings of the 79th Annual Convention of the American Psychological Association, 1971, 201-202.

Hobbs, E. D. Conservation of weight and volume in secondary school students. Unpublished M. Ed. theses, University of Alberta, 1972.

Hobbs, E. D. Adolescents' concepts of physical quantity. Developmental Psychology, 1973, 4, 431.

Hobbs, E. D. Methodological problems in conservation testing with particular reference to volume conservation. Alberta Journal of Educational Research, 1975, 21, (In Press).

Howe, A. C. The effect of science instruction based on task analysis on children's acquisition of the ability to perform formal operational tasks concerning volume relationships. (Doctoral dissertation, University of Texas of Austin.) Ann Arbor, Mich.: University Microfilms, 1969. No. 70-10,813.

Howe, A. Formal operational thought and the high school science curriculum. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Chicago, April, 1974.

Hyde, D. M. G. Piaget and Conceptual development. London: Holt, Rinehart, & Winston, 1970.

Inhelder, B. Criteria of the stages in mental development. In Tanner, J. M., & Inhelder, B. (Eds.). Discussion on Child Development. Vol. 1. London: Tavistock Publications Limited, 1956.

Inhelder, B. The diagnosis of reasoning in the mentally retarded. New York: John Day, 1968. Trans. from the French Le diagnostic du raisonnement chez les débiles mentaux. (2nd ed.) Neuchâtel: Delachaux et Niestlé, 1963.

Inhelder, B. Some aspects of Piaget's genetic approach to cognition. In Society for Research in Child Development. Cognitive development in children--five monographs of the Society for Research in Child Development. Chicago: The University of Chicago Press, 1970.

Inhelder, B. N., Bovet, A., Sinclair, H., & Smock, C. D. On cognitive development. American Psychologist, 1966, 21, 160-164.

Inhelder, B. N., & Piaget, J. The growth of logical thinking from childhood to adolescence. New York: Basic Books, 1958. Trans. from the French De la logique de l'enfant à la logique de l'adolescent. Vendôme (France): Presses Universitaires de France, 1955.

Inhelder, B. N., & Piaget, J. De la logique de l'enfant à la logique de l'adolescent. (2nd ed.) Vendôme (France): Presses Universitaires de France, 1970.

Inhelder, B., & Sinclair, H. Learning cognitive structures. In Massen, P., Langer, J., & Covington, M. (eds.). Trends and issues in developmental psychology. New York: Holt, Rinehart, & Winston, 1969.

Jackson, S. The growth of logical thinking in normal and subnormal children. British Journal of Educational Psychology, 1965, 35, 255-258.

Johnson-Laird, P. N., & Tagart, J. How implication is understood. American Journal of Psychology, 1969, 82, 367-373.

Keats, J. A. Formal and concrete thought processes. Princeton, N. J.: Educational Testing Services, 1955.

Kincheloe, J. W. The construction and validation of a formal thought instrument. (Doctoral dissertation, University of Northern Colorado.) Ann Arbor, Mich.: University Microfilms, 1972. No. 73-284.

King, W. H. Symposium: studies of children's scientific concepts and interests. I. The development of scientific concepts in children. British Journal of Educational Psychology, 1961, 31, 1-20.

Kuhn, D. Inducing development experimentally: comments on a research paradigm. Developmental Psychology, 1974, 10, 590-600.

Laurendau, M., & Pinard, A. Causal thinking in the child. Montreal: International Universities Press, 1962.

Laurendau, M., & Pinard, A. Reflexions sur l'apprentissage des structures logiques. In Bresson, F., & de Montmollin, M. (Eds.) Psychologie et épistémologie génétiques thèmes Piagétiens. Paris: Dunod, 1966.

Lawson, A. E. Relationships between concrete and formal operational science subject matter and the intellectual level of the learner. (Doctoral dissertation, University of Oklahoma.) Ann Arbor, Mich.: University Microfilms, 1973. No. 73-31,481.

Lawson, A. E., & Renner, J. W. A quantitative analysis of responses to Piagetian tasks and its implications for curriculum. Science Education, 1974, 58, 454-559.

Leon, L. O. The principle of conservation of invariance and its relationship to achievement in science in the junior high school. (Doctoral dissertation, University of Florida.) Ann Arbor, Mich.: University Microfilms, 1972. No. 72-16,626.

Lefrancois, G. R. The acquisition of concepts of conservation. Unpublished doctoral dissertation, University of Alberta, 1966.

Lefrancois, G. R. Jean Piaget's developmental model: equilibration-through-adaptation. Alberta Journal of Educational Research, 1967, 13, 161-171.

Lovell, K. The growth of basic mathematical and scientific concepts in children. London: University of London Press, 1961.

Lovell, K. Developmental processes in thought. Journal of Experimental Education, 1968, 37, 15-21.

Lovell, K., & Ogilvie, E. The growth of the concept of volume in junior school children. Journal of Child Psychology and Psychiatry, 1960, 1, 191-202.

Lunzer, E. A. Problems of formal reasoning in test situations. In Mussen, P. (Ed.) European research in cognitive development. Monographs of the Society for Research in Child Development, 1965, 30(2), 47-57.

Matulis, R. S. A survey of the understandings of selected concepts of logic by 8-18-year-old students. (Doctoral dissertation, University of Florida.) Ann Arbor, Mich.: University Microfilms, 1969. No. 70-14,901.

Mermelstein, E., & Shulman, L. S. Lack of formal schooling and the acquisition of conservation. Child Development, 1967, 38, 39-52.

Mecke, G., & Mecke, V. The development of formal thought as shown by explanations of the oscillations of a pendulum: a replication study. Adolescence, 1971, 6, 219-228.

Mischel, T. (Ed.). Cognitive development and epistemology. New York: Academic Press Inc., 1971.

Mussen, P., Langer, J., & Covington, H. (Eds.). Trends and issues in developmental psychology. New York: Holt, Rinehart & Winston, 1969.

Neimark, E. D. A preliminary search for formal operations structures. Journal of Genetic Psychology, 1970, 116, 223-232.

Nous, A. P. The effects of a Piaget-based program designed to induce in elementary and junior high school students basic correlational analysis using biological phenomena. (Doctoral dissertation, State University of New York at Buffalo.) Ann Arbor, Mich.: University Microfilms, 1970. No. 70-20,038.

Papalia, D. E. The status of several conservation abilities across the life-span. Human Development, 1972, 15, 229-243.

Pearce, E. E. Curriculum design for junior life sciences based on the theories of Piaget and Skinner. (Doctoral dissertation, University of Houston.) Ann Arbor, Mich.: University Microfilms, 1971. No. 72-012,911.

Peel, E. A. The pupil's thinking. London: Oldbourne, 1967.

Peluffo, N. Culture and cognitive problems. International Journal of Psychology, 1967, 2, 187-198.

Phillips, D. G. The development of the concept of displacement volume: a hierarchical model and its partial testing under two methods of presentation. Journal of Research in Science Teaching, 1971, 8, 9-19.

Piaget, J. The science of education and the psychology of the child. New York: Orion Press, 1970.

Piaget, J. The psychology of intelligence. Littlefield, Adams & Co., 1950. Trans. from the French La psychologie de l'intelligence. Paris: Colin, 1947.

Piaget, J. The child's conception of number. London: Routledge and Kegan Paul, 1952. Trans. from the French La genèse du nombre chez l'enfant. Neuchâtel: Delachaux et Niestlé, 1941.

Piaget, J. Logic and psychology. New York: Basic Books, 1957.

Piaget, J. Development and learning. In Ripple, R. E., & Rockcastle, V. (Eds.). Piaget rediscovered. Ithaca, New York: Cornell University, 1964a.

Piaget, J. Mother structures and the notion of number. In Ripple, R. E., & Rockcastle, V. (Eds.). Piaget rediscovered. Ithaca, New York: Cornell University, 1964b.

Piaget, J. Cognitive development in the child. Journal of Research in Science Teaching, 1964, 2, 176-186.

Piaget, J. Untitled article. Contemporary Psychology, 1967, 12, 532-533.

Piaget, J. Piaget's theory. In Mussen, P. H. (Ed.). Carmichael's Manual of Child Psychology (3rd ed., Vol. I). New York: John Wiley & Sons, Inc., 1970.

Piaget, J. The theory of stages in cognitive development. In Green, D. R., Ford, M. P., & Flamer, G. P. (Eds.). Measurement and Piaget. New York: McGraw-Hill, 1971.

Piaget, J. Intellectual evolution from aolescence to adulthood. Human Development, 1972, 15, 1-12.

Piaget, J., & Inhelder, B. N. The child's conception of geometry. London: Routledge & Kegan Paul, 1967. Trans. from the French La representation de l'espace chez l'enfant.

Piaget, J., & Inhelder, N. Le développement des quantites physiques chez l'enfant. (3rd ed.). Neuchâtel: Delachaux et Niestlé, 1968. (1st ed. 1941).

Piaget, J., & Inhelder, B. N. Mental imagery in the child. London: Routledge & Kegan Paul, 1971. Trans. from the French L'image mentale chez l'enfant. Vendôme (France): Presses Universitaires de France, 1966.

Piaget, J., & Inhelder, B. Memory and intelligence. London: Routledge & Kegan Paul, 1973. Trans. from the French Memoire et intelligence. Vendôme (France): Presses Universitaires de France, 1968.

Piaget, J., Inhelder, B. N., & Szeminska, A. The child's conception of geometry. London: Routledge & Kegan Paul, 1960. Trans. from the French La géométrie spontanée de l'enfant. 1947.

Pinard, A., & Laurendeau, M. "Stage" in Piaget's cognitive-developmental theory: exegesis of a concept. In Elkind, D., & Flavell, J. H. (Eds.) Studies in cognitive development: essays in honor of Jean Piaget. New York: Oxford University Press, 1969.

Reese, H. W., & Schack, M. L. Comment on Brainerd's criteria for cognitive structures. Psychological Bulletin, 1974, 81, 67-69.

Renner, J. W., Bibens, R. F., & Shepherd, G. D. Guiding learning in the secondary schools. New York: Harper & Row, 1972.

Renner, J., & Stafford, D. Teaching science in the secondary school. New York: Harper & Row, 1972.

Renner, J. W., Stafford, D. G., & Ragan, W. B. Teaching Science in the elementary school. (2nd ed.) New York: Harper & Row, 1973.

Ripple, R. E., & Rockcastle, V. N. (Eds.). Piaget Rediscovered. Ithaca, New York: Cornell University, 1964.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Siegler, R. Liebert, D., & Libert, R. Inhelder and Piaget's pendulum problem: teaching preadolescents to act as scientists. Developmental Psychology, 1975, 9, 97-101.

Sigel, I. E. The Piagetian system and the world of education. In Elkind, D., & Flavell, J. H. (Eds.) Studies in cognitive development: essays in honor of Jean Piaget. New York: Oxford University Press, 1969.

Sigel, I. E., & Hooper, F. H. (Eds.). Logical thinking in children. New York: Holt, Rinehart & Winston, 1968.

Somerville, S. C. The pendulum problem: patterns of performance defining developmental stages. British Journal of Educational Psychology, 1974, 44, 266-281.

Stephens, B., McLaughlin, J. A., & Mahaney, E. J. Ages at which Piagetian concepts are achieved. Proceedings of the 79th Annual Convention of the American Psychological Association, 1971, 203-204.

Stephens, B., McLaughlin, J. A., Miller, C. K., & Glass, G. V. Factorial structure of selected psycho-educational measures and Piagetian reasoning assessments. Developmental Psychology, 1972, 6, 343-348.

Stone, M. A., & Ausubel, D. P. The intersituational generality of formal thought. The Journal of Genetic Psychology, 1969, 115, 169-180.

Sullivan, E. V. Piaget and the school curriculum--a critical appraisal. Toronto: The Ontario Institute for Studies in Education, 1967.

Sullivan, E. V. Piagetian theory in the educational milieu: a critical appraisal. Canadian Journal of Behavioural Science, 1969, 1, 129-155.

Szeminska, A. The evolution of thought: some applications of research findings to educational practice. In P. H. Mussen (Ed.), European Research in Cognitive development. Monographs of the Society for Research in Child Development, 1965, 30(2), 47-57.

Taba, H. Curriculum development theory and practice. New York: Harcourt, Brace, & World, Inc., 1962.

Tanner, J. M., & Inhelder, B. (Eds.). Discussions on child development. Vols. I-III. London: Tavistock, 1958.

Teates, T. G. A comparison of the performance of ISCS and non-ISCS ninth-grade science students on several Piaget-type tasks. (Doctoral dissertation, Florida State University). Ann Arbor, Mich.: University Microfilms, 1970. No. 71-7113.

Tomlinson-Keasey, C. The nature of formal operations in preadolescence, adolescence, and middle age. (Doctoral dissertation, University of California, Berkeley.) Ann Arbor, Mich.: University Microfilms, 1970. No. 71-15,811.

Tomlinson-Keasey, C. Formal operations in females from eleven to fifty-four years of age. Developmental Psychology, 1972, 6, 364.

Towler, J. O., & Wheatley, G. Conservation concepts in college students: a replication and critique. Journal of Genetic Psychology, 1971, 118, 265-270.

Uzigiris, I. C. Situational generality of conservation. Child Development, 1964, 35, 831-841.

Verrizzo, O. Conceptions of conservation and reversibility in children of very superior intelligence. School Science and Mathematics, 1970, 70, 31-36.

Wason, P. C. Reason about a rule. Quarterly Journal of Experimental Psychology, 1968, 20, 273-281.

Weybright, L. D. Development and methodological issues in the growth of logical thinking in adolescence. (Doctoral dissertation, University of Illinois at Urbana-Champaign.) Ann Arbor, Mich.: University Microfilms, 1972. No. 72-19,957.

Wheeler, A. E. Student misconceptions in chemical equilibrium as related to achievement and cognitive level. Unpublished M.Ed. thesis, University of Alberta, 1973.

Wheeler, A. E., & Kass. H. Student misconceptions in chemical equilibrium as related to cognitive level and achievement. Paper presented at the 47th annual meeting of the National Association for Research in Science Teaching, Chicago, April, 1974.

Wheeler, D. Studies in the development of reasoning in school children. British Journal of Statistical Psychology, 1958, 11, 137-159.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Wood. D. A. The Piaget-process matrix, School Science and Mathematics, 1974, 74, 407-412.

Yudin, L. W. Formal thought in adolescence as a function of intelligence. Child Development, 1966, 37,697798 697-798.

Yudin, L., & Kates, S. L. Concept attainment and adolescent development. Journal of Educational Psychology, 1963, 54, 177-182.

APPENDIX A

TESTS

Volume Conservation

The response sheet for the volume conservation task is reproduced as Figure 2. The instructions for the tester are as follows:

1. Say "Write your name, sex, and birthday at the top of the answer sheet. After each question, don't go back and change anything."
2. Place three glasses on a table in front of you so that all subjects can see. Fill each glass to the mark. Draw subjects' attention to the fact that the level is the same in each glass.
3. Show subjects the six plasticene balls which are the same size and point out that they are the same size.
4. Say "What do you think will happen if I place this ball in this [middle] glass? I want to know if it will take up any room in the water. Where will the water come to? Lets try it."
5. Place the ball into the middle glass and say "See? The water came up to here."
6. Say "Now if I put this ball [one the same as the first] into this [left-hand] glass, I want to know

how much room it will take up in the water. Where will the water come to? You can show me by putting a line on the third glass in the answer sheet. Put it beside number one. Draw a line on the third glass to show where you think the water will come to. I'll show you. If you think the water comes to here [indicating a level on the left-hand glass], draw a line like this [demonstrate on an answer sheet or a blackboard]. If you think the water only comes to here, put the line like this. Do it now." Pause, then ask "Is anybody not sure what to do?" If any subject is not sure, repeat the explanations once.

7. Say "Let's try it." Place the ball in the glass on your left. Say "See? the water comes to here--the same as in this [middle glass]. If you weren't correct, please don't go back and change your answer." Remove the ball from the glass on your left.

8. Take the small plasticene ball. Say "Now let's take this small ball. I want to know how much room it will take up in the water. If I put <u>this</u> ball in the water, where will the water come to? Mark it beside number two on the answer sheet. Do it now. Number two."

9. Wait, then say "OK let's try it." Place the ball in the glass on your left and say "See? The water came up to here."

10. Repeat 8 and 9 three times substituting first, the large plasticene ball (answer sheet number three); second, the steel ball the same size as the six standard balls (answer sheet number four); and third, the small lead ball (answer sheet number five).
11. Say "Now let's do something else. We start with these two the same size. [Indicate two of the standard balls.] We take this one and roll it out into a sausage. If I put the ball in the water, the level comes up to here [indicating the level in the middle glass]. What if I put the sausage into the water? I want to know how much room the sausage will take up in the water. Where will the water come to? Mark it beside number six." Do not demonstrate immersion this time.
12. Repeat step 11 twice, first breaking a ball up into five approximately equal pieces; and second squashing a ball into a "pancake."

CV-A

The question and answer sheet for CV-A is reproduced on page 121. The instructions for the tester were as follows:

1. Give out the answer sheets. Caution subjects not to say anything, just to write. Have them write

their names on the sheets.

2. Hold up the brass and aluminum rods, one in each hand, so that the rods are vertical, about 50 cm apart, the brass rod on your left. Say "These are the rods you are supposed to look at." Allow five seconds, then hold them horizontally, side-by-side, about 3 cm apart, the aluminum rod on top aligned at neither end. Allow five seconds, then place the rods about 3 cm apart on a desk or table so that they stand vertically with the brass rod on your right.
3. Read the test to the subjects saying "Question one says. . . ."
4. Allow 10 seconds for subjects to write answers to question one and two minutes for question two. Stop the test 20 minutes after you begin.

Note: if subjects and questions related to clarifying the test, say "Do what you think is right."

CV-B

The question and answer sheets for CV-B are reproduced on pages 112-113. The instructions for the tester were as follows:

1. Give out the kits of wood and the tests.
2. Have any new subjects write their name, sex, and birthday at the top of the answer sheets. Other subjects write only their name.

3. Have subjects line up their planks on their desks in alphabetical order with the left-hand edges aligned. Say "You can move the planks around all you want when the test starts, but they should be like that to begin with.

4. Say "See that some are pine and some are redwood. Have the markings up. Now when we say 'thickness' we mean this way [indicate] from top to bottom. That's thickness. This way [indicate] is width. This way [indicate] is length. Put your finger on a wide, thick plank." Check that each subject is pointing to a correct plank. Say "Put your finger on a narrow, thick plank." Check again. Explain the difference again to any subjects whose responses indicate confusion between length and thickness.

5. Say "Now let's read the test." Slowly read through the entire test.

6. Say "Now it's very important that you say which planks you would use to do the tests. Name the planks by their letters, for instance you can say 'plank A', or 'plank F' and so on.

7. Tell subjects to begin. Allow 20 minutes for the entire administration.

<u>CS</u>

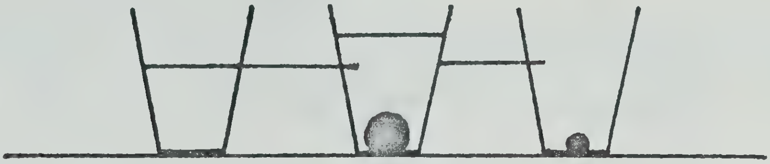
The question and answer sheets for CS are reproduced as Figures 6 and 7. The instructions for the tester were as follows:

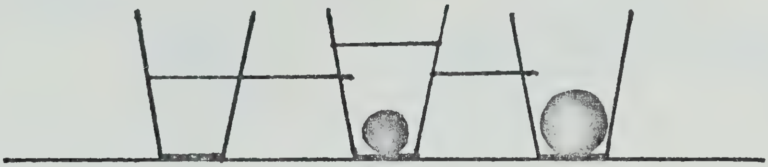
1. Say "You'll have 20 minutes to try this little test. Look at it, and let's read it together."
2. Read it slowly with added explanation as follows:
   (a) after "radio" in line 1 say "like that little picture.
   (b) say "for" before the price each time.
   (c) after "$10" say "that's those crisscross lines."
   (d) after "$2" say "That thing at the side."
   (e) after "$1" pause five seconds.
   (f) after "draw" say "whichever you like."
   (g) after "$45" say "Let me show you another one. You could have the strap and the battery [draw on chalkboard] for $40 plus $5 plus $2. That is $47."
   (h) after "sheets" ask "Is anybody not sure what to do?"
3. Answer any questions related to what to do, how to write answers, etc. Record such questions. Do not answer questions related to possible number of combinations. Say "I want you to figure it out,"

or "I won't tell you that, that's part of the test."

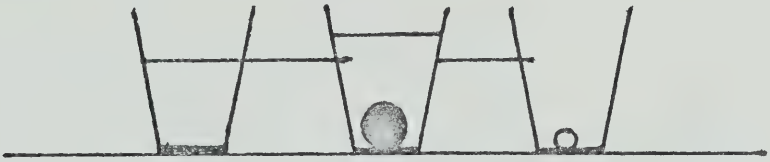
4. Allow 15 minutes for subjects to respond.

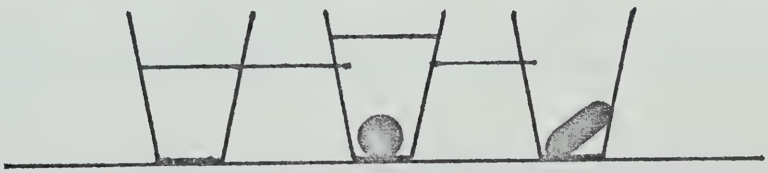
1. 

2. 

3. 

4. 

5. 

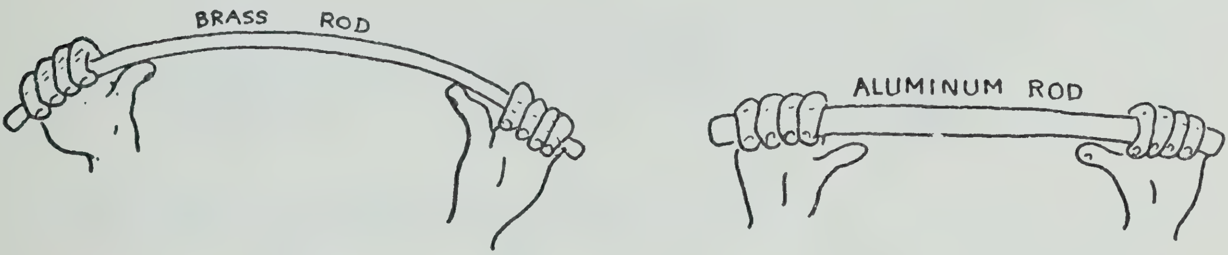
6. 

7. 

8. 

TEST NUMBER TWO

Look at the two metal bars being held up by the tester.
The silvery one is made of ALUMINUM.
The other one is made of BRASS.
A student wanted to find out if rods made of brass bend more easily than rods made of aluminum. The student tried to bend each rod. The pictures show what happened.



QUESTIONS:

1. Does this prove that rods made of brass bend more easily than rods made of aluminum? ______________

2. Why or why not? ______________

3. Give two ways you could make this a better test to see if rods made of brass bend more easily than rods made of aluminum.

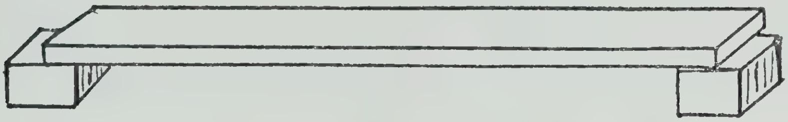
   (1) ______________

   (2) ______________

TEST NUMBER FOUR

How to test the stiffness of a wooden plank:—

A wooden plank can be tested for stiffness by placing it on two blocks like this,



Then putting a weight on it like this:

The less it bends, the stiffer it is.

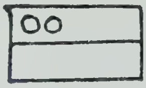
You have been given a number of planks, some wide, some narrow, some thick, some thin, some short, some long, some made of pine, some made of redwood.

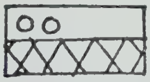
1. Suppose you want to find out if the length of a plank makes any difference to its stiffness. How would you find out? (say which planks you would use)

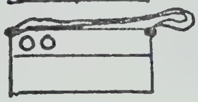
2. Suppose you want to find out if the thickness of a plank makes any difference to its stiffness. How would you do a test to find out? (say which planks you would use)

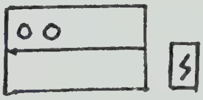
3. Suppose you wanted to test whether the kind of wood affects the stiffness of a plank (pine or redwood). How would you find out? (say which planks you would use)

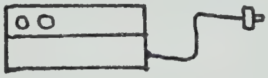
TEST NUMBER THREE

A transistor radio  costs $40, but you can buy any of the following "extras" if you want to:

leather case  $10.

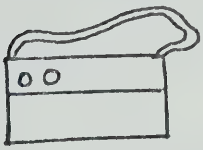
OR carrying strap  $5.

OR extra battery  $2.

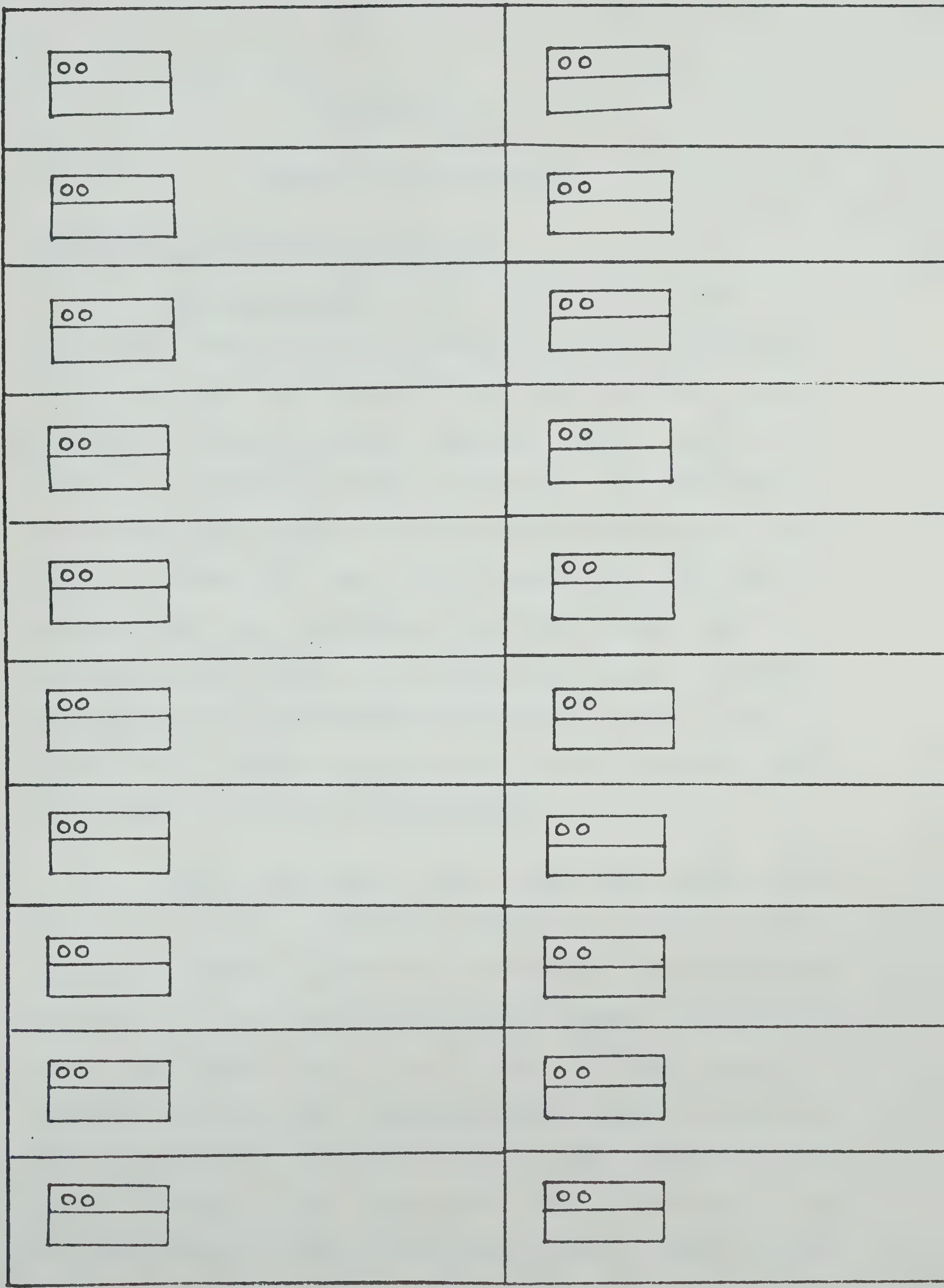
OR ear phone  $1.

Suppose you want to buy a radio. Write down or draw as many different choices as you can think of, and the prices. For example

 radio + case + earphone $51.

 radio + strap $45.

Put your answers on the special answer sheets



APPENDIX B

INSTRUCTIONAL SEQUENCES

Controlling Variables Lesson One

Sugar experiment. Say "I'm going to prove that sugar dissolves more quickly in a short glass than in a tall glass OK? Watch." Fill the tall glass about two-thirds full with water, the short glass almost full. Place a tablespoon of sugar in the tall glass and stir for about eight seconds so that not all dissolves. Then place a teaspoon of sugar in the short glass and stir rapidly until all dissolves. Say "that proves sugar dissolves more quickly in a short glass right?" Discuss the experiment with subjects asking how to make it a "good test." Establish that equal volumes of water, and equal amounts of sugar are required.

Rubber experiment. Say "I have some rubber bands here" and produce a number of red and blue rubber bands varying according to thickness of rubber, width of rubber, and length. Say "someone told me red rubber is 'stretchier' than blue rubber. Let's try it out. I'll put a weight [large wrench] on this red one [large, thick] and see how far it stretches. Now put a weight [small pliers] on this blue one [small, thin] and see how far it stretches." Hold the two weighted rubber bands side by side, then put down

the weights and hold up the rubber bands side by side. Say "So that proves it--red rubber is stretchier than blue rubber, right?" If subjects object that the bands are not the same size, repeat using bands the same length but different width. Continue, discussing with subjects, until the requirements of equal weights, and equal dimensions for the rubber bands are established.

Controlling Variables Lesson Two

Light bulbs. Place on a table between instructor and subjects the lamp base and the array of light bulbs (as shown in Figure 1 [p. 53]). Mount lamp B (24W clear picture bulb) in the base and switch on, then off. Demonstrate the photographic light meter. Indicate where the light enters and show where the needle registers. Switch on the picture light and show the needle reading. Switch off to show how the needle returns to zero. Say "This is a 25 watt light bulb. How can we find out if light bulbs with higher watts make the needle go up more? Through discussion, establish which pairs of bulbs provide a test of whether the wattage of the bulb affects the meter reading. Repeat the procedure to establish proper tests of whether meter reading is affected by shape of bulb, or frosting.

Paper strips. Say "Let's try something different." Give each student a package of paper strips. Say "Put these on your desk. There are two colours, two lengths, two widths. Pick out a pair the same colour and width, but with different lengths. [Check, help individuals where necessary.] Find another pair. Another. Another." Repeat this procedure, first for pairs differing only in width, then for pairs differing only in colour. Do not mention controlling variables or experimental tests.

Rubber tubes. Show subjects the array of rubber tubes differing in length, diameter, and colour (red and black). Ask individual subjects to select pairs of tubes "the same except for length" then "colour," then "thickness."

Metal rods. Repeat the rubber tubes procedure using the array of steel and copper rods of varying length and diameter.

Tasks for Today

The third lesson in the controlling variables sequence was, together with the third lesson in the combinatorial sequence, formed the classroom assignment entitled "Tasks for Today," which is reproduced on pages 132 to 137 . Questions 1, 2, and 3 formed lesson

three--the combinatorial sequence, while question 4 and 5 formed lesson three in the controlling variables sequence. Instructors were told to hand out the sheets to subjects and ask them to begin at question 1. Instructors were told to help subjects as requested, and to direct subjects to begin question 4 after 20 minutes irrespective of progress on questions 1, 2, and 3.

<u>Combinatorial Sequence Lesson One</u>

<u>Objective</u>. Subjects will understand the triangles problem: given a packet containing 40 cardboard triangles--10 blue, 10 black, 10 red, 10 yellow, place none, one, two, three, or four triangles in baking cups according to the following rules:

(a) no two baking cups have similar contents

(b) only one triangle of any one colour in any one baking cup.

Stated behaviorally: places triangles in baking cups correctly according to above rules.

<u>Equipment</u>. For each subject:

(a) one white envelope containing 10 red, 10 blue, 10 yellow, 10 black cardboard triangles;

(b) set of 20 baking cups clipped with paper slide.

<u>Method</u>. 1. Say "I'm going to get you to do a little task. You might find it very easy, you might not."

2. Give out the paper triangles, and say "There should be 40 triangles in your envelope. Empty them out on the desk. See that there are ten of each colour." Wait until everyone is sure of having the correct number of triangles.

3. Say "OK, now I'm going to give you some baking cups and then I'll explain the rules of the task." Caution students that cups are light and blow around easily, and tell them to put the cups out in rows on the desk.

4. Say "Here are the rules. You have to put none, one, two, three, or four triangles in a cup. For instance, a red and a blue, or just a red, or a black, a yellow and a blue. [Demonstrate.] You can't have more than one triangle of the same colour in any cup. [Demonstrate.] No two cups can have the same sets of triangles in. You have to make as many different sets as you can. [Demonstrate.] Any questions? Go ahead and make as many different sets as you can." The colours of the cups have nothing to do with the experiment.

5. Point out mistakes and explain as class works.

6. Get students to clip cups together and put triangles back in envelope.

Note: assist wherever necessary--this is instruction, not testing.

Combinatorial Sequence Lesson Two

1. Say "Today I will give you a way to do the task with the coloured triangles."
2. Give subjects each a packet of triangles and a pink sheet ruled into 20 squares.
3. With the aid of four large triangles, show subjects how to generate the combinations by taking one at a time, two at a time, three at a time, and four at a time.
4. Let subjects try it once with the paper triangles, assisting individuals where necessary. Each combination is to be placed in a square on the pink sheet.
5. Take back the packets of triangles and give subjects packets containing 10 plastic counters ("bingo chips"), 10 paper slides, 10 paper fasteners, and 10 pennies. Ask them to do the same thing with these objects instead of the triangles.
6. Have students return the packets of objects, but allow them to keep the pennies.

TASKS FOR TODAY

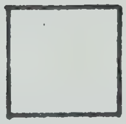
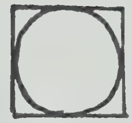
1. Make as many combinations as possible with the set of 4 kinds of objects in the packet supplied. Use the pink sheet ruled in squares to put the combinations on. Ask for assistance if you need it.

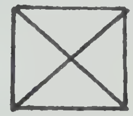
2. When you have finished number 1, pick one of the following and do it:--

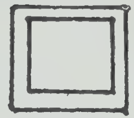
   A. CLOTHES. Suppose you have bought:
   (1) a new belt
   (2) a new hat
   (3) a new sweater
   (4) a new pair of sunglasses
   Figure out all the ways you could wear these. For example you could wear the belt and hat, or just the sweater or all four. Use the pink sheets to write down the combinations (or draw them).

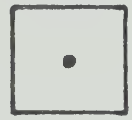
   B. HAMBURGERS. With a hamburger you can get:
   (1) ketchup
   (2) relish
   (3) cheese
   (4) onions
   Figure out all the different ways you could have the hamburger. Use the pink sheet to put down your answer.

   C. CARS. When you buy a new car you can get, as extras, any of these:
   (1) whitewall tires
   (2) automatic transmission
   (3) radio
   (4) heater
   Figure out the different ways you can buy the car. For example you could get the whitewalls and the radio, or just automatic transmission, or all four. Use the pink sheet.

3. A square () can have a circle in it 

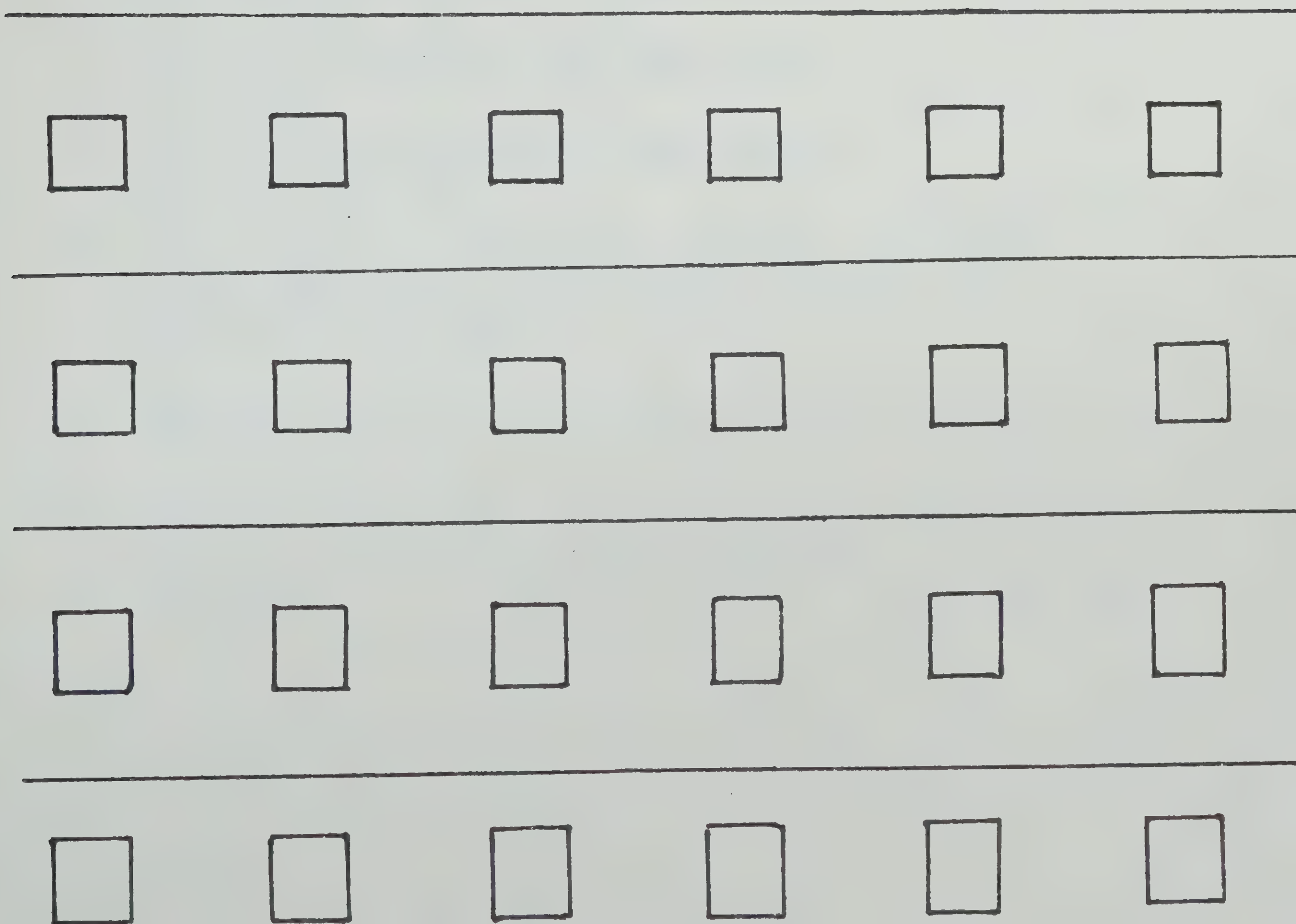
a cross in it 

a border on it 

or a dot in the middle 

Work out all the different ways you can have a square.
For example:

 or  and so on.



4. Look at the sheet with pictures of screws, buttons, and cups on it. Try these:

A. The screws are different lengths and thickness, and have different shaped heads.

Pick out two that are the same except for length: ____ and ____

Pick out two that are the same except for shape of the head: ____ and ____

Pick out two that are the same except for thickness: ____ and ____

Another two: ____ and ____

B. The buttons are different in the following ways, some have rims, some don't. Some have two holes, some four. Some are large, some small.

Pick out two that are the same except for different rims: ____ and ____

Pick out two that are the same except for size: ____ and ____

Pick out two that are the same except for the number of holes: ____ and ____

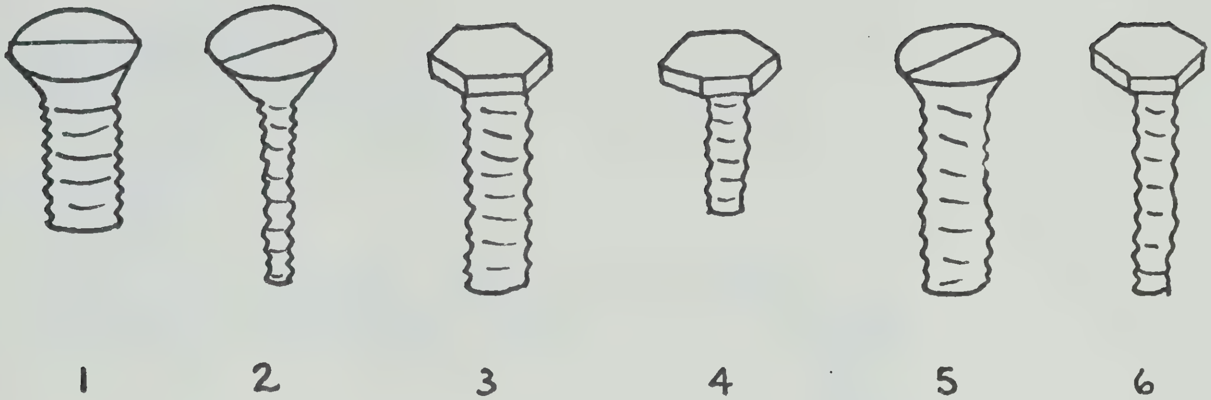
C. The glasses are different as follows: some have a wide base, some a narrow base. Some are tall, some are short. Some are "straight-sided" while others are "tapered."

Pick out two glasses different only in height: ____ and ____

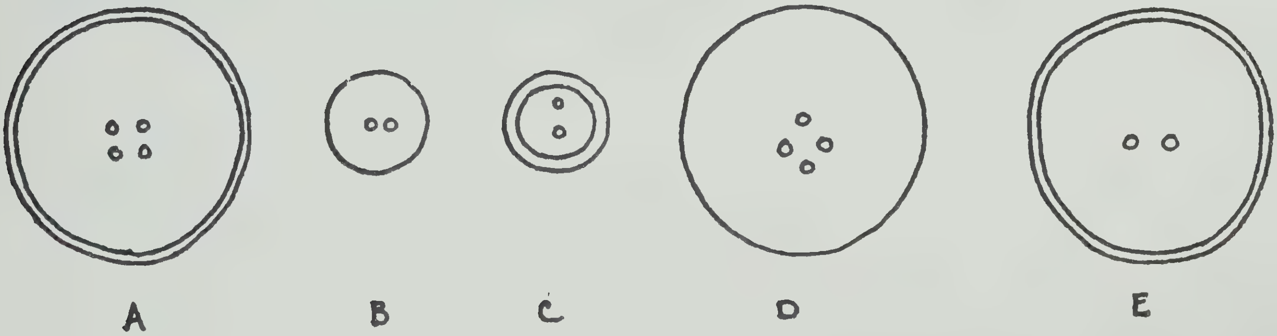
Pick out two different only in shape ____ and ____

Pick out two different only in size of the base: ____ and ____

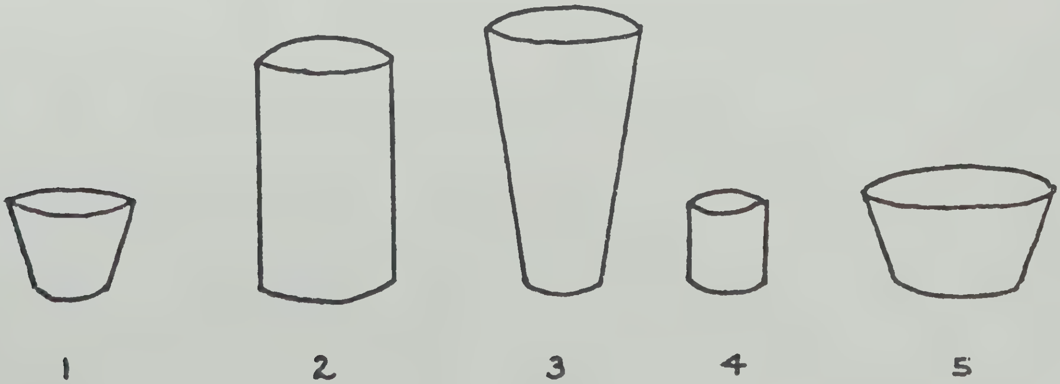
Screws



Buttons



Glasses



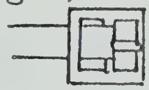
5. The belts problem

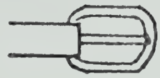
The belts differ as follows:

_____ in length

_____ in width

_____ in shape at the end (pointed or straight)

_____ in shape of the buckle (square 

or round )

I. Pick out two belts the same except for width: _____ and _____

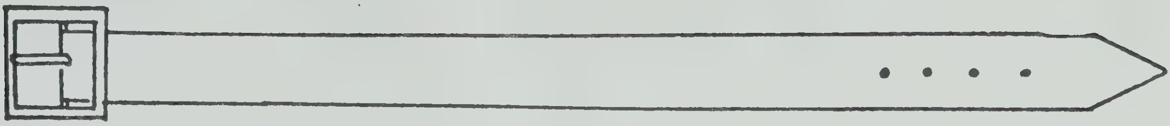
II. Pick out two the same except for type of buckle: _____ and _____

III. Pick out two the same except for shape at the end: _____ and _____

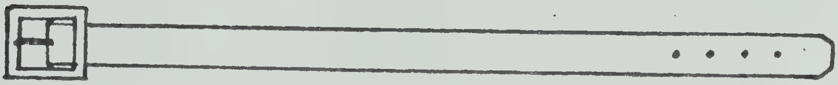
IV. Pick out two the same except for length: _____ and _____

Belts

A



B



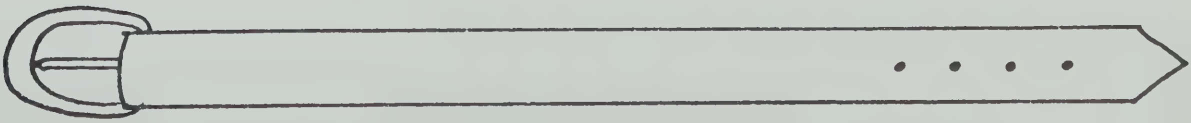
C

D

E

F

G



H